Heavy Traffic Analysis of the Dynamic Stochastic Inventory-Routing Problem

Martin I. Reiman, Rodrigo Rubio, and
Lawrence M. Wein

#3881-96-MSA February, 1996

Heavy Traffic Analysis of the Dynamic Stochastic Inventory-Routing Problem

Martin I. Reiman
AT&T Bell Laboratories
Murray Hill, NJ 07974

Rodrigo Rubio
Operations Research Center, M.I.T.
Cambridge, MA 02139

Lawrence M. Wein
Sloan School of Management, M.I.T.
Cambridge, MA 02139

ABSTRACT

We analyze three queueing control problems that model a dynamic stochastic distribution system, where a single capacitated vehicle serves a finite number of retailers in a make-to-stock fashion. The objective in each of these vehicle routing and inventory problems is to minimize the long run average inventory (holding and backordering) and transportation cost. In all three problems, the controller dynamically specifies whether a vehicle at the warehouse should idle or embark with a full load. In the first problem, the vehicle must travel along a prespecified (TSP) tour of all retailers, and the controller dynamically decides how many units to deliver to each retailer. In the second problem, the vehicle delivers an entire load to one retailer (direct shipping) and the controller decides which retailer to visit next. The third problem allows the additional dynamic choice between the TSP and direct shipping options. By assuming that the system operates under heavy traffic conditions, we approximate these queueing control problems by diffusion control problems, which are explicitly solved in the fixed route problems, and numerically solved in the dynamic routing case. Simulation experiments confirm that the heavy traffic approximations are quite accurate over a broad range of problem parameters. Our results lead to some new observations about the behavior of this complex system.

February 1996

1 Introduction

A prototypical example of the inventory-routing problem (IRP) is the challenge faced by a large oil company as it distributes gasoline to its various gas stations: several warehouses, which hold inventory of a particular item (gasoline), serve a set of retailers (stations) in a make-to-stock fashion; arriving customers (automobiles) consume the product at these retail sites, and a fleet of finite capacity vehicles (tanker trucks) is used to transport the product from the warehouse to the various retailers.

The management decisions involved in the design and operation of such a system are many-fold and complex. Traditionally, a hierarchical decomposition of the problem is used to allow for a solvable model at each of the levels (e.g., Simchi-Levi 1992). At the strategic level, the managers of this system must determine the location and number of warehouses and retailers, as well as the assignment of retailers to warehouses. At a tactical level, they must decide on the number of vehicles to operate, and possibly on the assignment of vehicles to service districts. At the operational level the decisions include: whether to send a particular vehicle out or let it idle, how much of the capacity of the vehicle to use, which of the retailers should each vehicle visit, and how much of its load should a vehicle deliver to each of the retailers on its route.

At the tactical and operational levels, the essence of the IRP is the tradeoff between inventory costs and transportation costs: in order to reduce inventory levels at the retail sites without affecting the service level, more frequent replenishment deliveries are required, thereby increasing the transportation costs. In many applications, customer demand (and to a lesser extent, vehicle travel times) is subject to considerable stochastic variation. In such cases, a stochastic model is required to accurately capture the inventory costs, and in this paper we focus on the operational aspects of the IRP in a dynamic stochastic setting.

The field of operations research is sometimes criticized because real-world applications have lagged behind theoretical progress (e.g., Ackoff 1987). The IRP is an important counterexample to this perception: while heuristics for this notoriously difficult problem have led to some spectacularly successful industrial applications at both the operational (e.g., the

Edelman Prize winning work of Bell et al. 1983, Golden et al. 1984) and tactical (Larson 1988) levels, a concomitant mathematical theory for the IRP in a dynamic and stochastic environment has not been forthcoming. Federgruen and Zipkin (1984) analyze a single-period IRP with stochastic demand, and Dror and Ball (1987) develop a heuristic technique to reduce the long run average problem to a single period problem. Recent studies that consider the operational aspects of the stochastic IRP include Trudeau and Dror (1992), who develop heuristics for the case of an external supplier, where retailer inventories are only observable at delivery times; Minkoff (1993), who constructs a decomposition heuristic for a Markov decision model that dispatches vehicles on a prespecified set of itineraries, where each itinerary is characterized by an inventory allocation to a subset of customers; and Kumar, Schwarz and Ward (1995), who develop myopic static and dynamic strategies for allocating the contents of a vehicle to the various retailers on a predetermined tour. Chan, Federgruen and Simchi-Levi's (1994) probabilistic analysis of random instances of the deterministic IRP is useful for addressing tactical and strategic issues, but has no bearing on the operational aspects of the IRP with stochastic demand.

Our system model has one warehouse and one capacitated vehicle; hence, we effectively assume that the higher level decisions have been made to assign a single warehouse and a single vehicle to serve all retailers in a particular region. An ample amount of inventory is available at the warehouse, and the cost of holding this inventory is not included in the model. Retailer demand and vehicle travel times are random, unsatisfied demand is backordered, and the objective is to minimize costs due to holding and backordering inventory (cost rates may differ by retailer) and operating the vehicle. One of the crucial decisions in our problem is the *vehicle idling policy*: when the vehicle is at the warehouse, the controller can either send the vehicle out with a full load or let the vehicle sit idle; because we employ a long run average cost criterion, if the vehicle does not idle and the "traffic intensity" of the system is less than one, then an infinite amount of retailer inventory will build up over the long run.

Two types of IRP's are analyzed: the first assumes fixed routing and the second allows dynamic routing. We consider two variants of the fixed routing IRP: in the IRP with *TSP routing*, when a vehicle leaves the warehouse it uses a pre-optimized tour of the m retailers,

which we refer to as the TSP (travelling salesman problem) tour. In addition to the vehicle idling policy, the controller must decide how many units to deliver to each retailer, and this decision is based on the current inventory levels at all retailers and on the remaining number of units in the vehicle. The second variant is the IRP with *direct shipping*; in this case, each time the vehicle leaves the warehouse it delivers all of its contents to a single retailer, and the controller dynamically specifies which retailer to visit next. In the IRP with *dynamic routing*, the controller decides, based upon the current inventory levels, whether to use a TSP tour or direct shipping.

By only allowing TSP tours or direct shipping, we avoid an assault on the combinatorial aspects of the embedded routing problem, and model these problems as queueing control problems. Since these control problems appear to be analytically intractable, heavy traffic analysis is employed in order to make further progress. To obtain an interesting and nontrivial limiting control problem, we assume that the server (vehicle) must be busy most of the time to meet average demand, the vehicle capacity is large, the tour completion time is long and nearly deterministic, and the vehicle operating cost is large; these heavy traffic conditions are stated more precisely later in the paper. Guided by the heavy traffic limit theorems in Coffman, Puhalskii and Reiman (1995a,b), we uncover a time scale decomposition in the heavy traffic limit; however, no weak convergence proofs are provided because they would be very demanding and would distract us from our primary objectives: (i) to gain insights into the nature of the optimal policies for the IRP, and (ii) to develop effective policies for the operation of these systems. Under the traditional heavy traffic normalization, where time and inventory are compressed by a factor of n and $\sqrt{n}$, respectively (and n is going to infinity), the (embedded or averaged) total retailer inventory process is well approximated by a one-dimensional reflected Brownian motion on the interval $(-\infty, w]$, where the control parameter w represents an aggregate base stock level that dictates the vehicle idling policy. The drift and variance of the Brownian motion depend on whether the TSP or direct shipping policy is being used, and in the dynamic routing problem the controller dynamically switches between two Brownian motions (one for TSP and one for direct shipping), each possessing its own drift, variance and cost structure. If we take the heavy traffic

normalization and slow down time by a factor of $\sqrt{n}$, then a fluid limit is obtained on this faster time scale. A deterministic analysis at this time scale allows us to optimize over the other operational decisions.

Our results are quite explicit; the derived controls for the fixed routing IRPs are either given in closed form or in terms of parameters that are solutions to certain equations. Because we cannot characterize the form of the optimal policy for the diffusion control problem associated with the IRP with dynamic routing, we analyze a class of triple threshold policies that is conjectured to contain the optimal policy, and numerically compute the optimal policy to the diffusion control problem. A computational study is carried out that confirms the accuracy of the heavy traffic analysis and allows us to obtain insights into the relative importance of the various operational decisions (e.g., the vehicle idling policy, static vs. dynamic allocation, TSP vs. direct shipping, fixed vs. dynamic routing).

In §2 and §3, we analyze the IRP with TSP routing and direct shipping, respectively. The performance of these two routing schemes is compared in §4 and the IRP with dynamic routing is analyzed in §5. Our computational study is described in §6. §7 contains some concluding remarks, including a summary of our key findings.

2 The IRP with TSP Routing

2.1 Problem Formulation

Consider a system where a single vehicle with capacity V is used to distribute a standard product to m geographically dispersed retailers. An infinite supply of the product is kept at the central warehouse at no cost. Customers are served from the retailer inventories in a make-to-stock fashion, and demand that cannot be served immediately is backordered. When the vehicle is operating the following policy is used: the vehicle leaves the warehouse (indexed as station 0) with a full load and then visits all the retailers in a predefined sequence before returning empty. Alternatively, the vehicle may idle at the depot. Though the order in which retailers are visited could be arbitrary, we assume that it is the solution to the

implied TSP, and refer henceforth to this service scheme as the TSP policy. Without loss of generality, we assume that retailers are indexed from 1 through m according to their position in the TSP tour.

Two sources of variability are considered: customer demand and travel times. For $i = 1, \ldots, m$, customer demand at retailer i occurs according to an independent renewal process $\{D_i(t), t \geq 0\}$ with rate λ_i and squared coefficient of variation c_{di}^2 (variance of the interdemand time divided by the square of the mean). The cumulative total demand in $[0, t]$ is denoted by $D(t) = \sum_i D_i(t)$, and $\lambda = \sum_i \lambda_i$ is the total demand rate. (In all summations of this paper the index runs over the set of retailers $\{1, 2, \ldots, m\}$, unless explicitly indicated otherwise.) Our results easily generalize to cases with correlated compound renewal processes; see §6 of Reiman (1984) for details. The sequence of travel times between facilities i and j is given by iid samples of the random variable T_{ij}, which has mean θ_{ij} and squared coefficient of variation c_{ij}^2 (i, j run from 0 to m). These travel times are independent of the demand streams and of each other. Keeping with the convention in the literature, we assume that pickup and delivery of units occur instantaneously; in practice, load/unload times tend to be dwarfed by the travel times. (Although non-zero load/unload times can be incorporated in a straightforward manner, the analysis becomes more tedious and its inclusion would cloud the basic issues). Hence, the mean and variance of the total time required to complete the TSP tour are given by $\theta_T = \sum_{j=0}^{m-1} \theta_{j,j+1} + \theta_{m0}$ and $s_T^2 = \sum_{j=0}^{m-1} \theta_{j,j+1}^2 c_{j,j+1}^2 + \theta_{m0}^2 c_{m0}^2$, respectively, where the subscript "T" is mnemonic for TSP. For later use, we define the squared coefficient of variation of the tour completion time as $c_T^2 = s_T^2/\theta_T^2$, and let $\{S_T(t), t \geq 0\}$ denote the counting process for TSP tour completions up to time t assuming the vehicle is continuously active in $[0, t]$.

Because the route is fixed, only two operating control decisions remain: (i) whether the vehicle should be busy or idle; (ii) while the vehicle is busy, how much of the load to leave at each retailer. The busy/idle control is expressed in terms of the cumulative process $B_T(t)$, which represents the amount of time the vehicle is busy in $[0, t]$. We do not allow tours to be interrupted, and so the sequence τ_k, $k = 1, 2, \ldots$ of tour completion epochs is given by $\tau_k = \inf\{t \mid S_T(B_T(t)) \geq k\}$. The delivery allocations are modeled by the m-dimensional

control process $L_i(t)$, which represents the cumulative amount delivered to retailer i up to time t. In anticipation of future developments, let us express this control in terms of a nominal delivery size for retailer i, denoted by V_i, and a dynamic allocation process $\epsilon_i^T(t)$. We let $V_i = \lambda_i V/\lambda$ for all i, so that the nominal delivery size corresponds to allocating the vehicle capacity V among the retailers according to their relative demands. The load allocation process is defined by

$$\epsilon_i^T(t) = L_i(t) - V_i S_T(B_T(t)) \quad \text{for} \quad t \geq 0, \tag{1}$$

which represents the cumulative deviations from the nominal delivery size over past tours, plus the amount delivered during the current cycle for retailer i. Because the tour completion history can be observed, we need only specify the value of $\epsilon_i^T(t)$ to determine the total deliveries to retailer i up to time t. Notice that deviations from the nominal allocation cancel out across the retailers and the process $\epsilon_T(t) = \sum_i \epsilon_i^T(t)$ represents the total amount delivered during the current cycle. Because we assume that the vehicle leaves the warehouse with a full load and returns empty, the dynamic load allocation process must satisfy

$$\begin{aligned}
\epsilon_i^T(0) &= 0 \quad \text{for all} \quad i, &(2)\\
\epsilon_i^T(t^+) &> \epsilon_i^T(t^-) \quad \text{only if retailer } i \text{ is visited at time } t, &(3)\\
\epsilon_i^T(t) &\geq \epsilon_i^T(\tau_{k-1}) \quad \text{for} \quad t \in (\tau_{k-1}, \tau_k) \quad \text{and all} \quad i, &(4)\\
\epsilon_T(\tau_k^-) &= V \quad \text{and} &(5)\\
\epsilon_T(\tau_k) &= 0, &(6)
\end{aligned}$$

where the superscripts "-" and "+" denote the times just before and after an epoch.

The number of units in inventory (or backordered if this quantity is negative) at retailer i at time t is denoted by $Q_i(t)$, and the total inventory at the retailers is $Q(t) = \sum_i Q_i(t)$. If we assume that $Q_i(0) = 0$ (which is without loss of generality, since a long run average cost criterion is being used) then the current inventory $Q_i(t)$ equals the cumulative deliveries minus the cumulative demand, which by (1) is given by

$$Q_i(t) = V_i S_T(B_T(t)) - D_i(t) + \epsilon_i^T(t) \quad \text{for} \quad i = 1, \ldots, m, \quad t \geq 0. \tag{7}$$

Define the cumulative vehicle idle time process $I(t)$ by

$$I(t) = t - B_T(t) \ \text{for} \ t \geq 0, \tag{8}$$

so that the control policy $B_T(t), \epsilon_i^T(t)$ must satisfy

$$\begin{aligned} B_T, \epsilon_i^T \quad & \text{are nonanticipating with respect to } Q, & (9) \\ B_T \quad & \text{is nondecreasing and continuous with } B_T(0) = 0, & (10) \\ I \quad & \text{is nondecreasing with } I(0) = 0. & (11) \end{aligned}$$

Our objective function includes transportation costs and inventory holding and back-ordering costs. The travel cost rate per unit time, which includes vehicle depreciation, fuel and driver cost, is r. Note that these costs can be combined because we are ignoring the load/unload times (only the driver, but not the vehicle, is busy while loading and unloading). Inventory costs are assumed to be piecewise-linear, with the holding cost rate (per unit in inventory per unit time) at retailer i denoted by h_i and the backorder cost rate by b_i. Because travel costs are incurred whenever the vehicle is busy, the travel cost rate r can be equivalently treated as a *reward* for exerting idleness. Hence the problem reduces to finding a control policy $\left(B_T(t), \epsilon_i^T(t)\right)$ to minimize

$$\limsup_{T \to \infty} \frac{1}{T} E\left[\int_0^T \sum_i (h_i\{Q_i(t)\}^+ + b_i\{Q_i(t)\}^-)\, dt - rI(T)\right] \tag{12}$$

subject to (2) - (11), where the "+" and "-" denote the positive and negative parts.

The dynamic stochastic IRP, as formulated in (2) - (12), does not seem to be tractable. Even under Markovian assumptions for the underlying random processes, the control space is enormous and the state space has $m+2$ dimensions: the inventory/backorder level at each retailer and the location and total contents of the vehicle. To gain further understanding of the problem, we analyze it when the system operates in the heavy traffic regime.

2.2 Heavy Traffic Normalizations

We begin our heavy traffic development by centering the service completion and demand processes; define the centered processes $\mathcal{S}_T(t) = S_T(t) - \theta_T^{-1}t$ and $\mathcal{D}(t) = D(t) - \lambda t$. It is

convenient to define the process

$$\chi(t) = \left(\frac{V}{\theta_T} - \lambda\right)t + V\mathcal{S}_T(B_T(t)) - \mathcal{D}(t). \tag{13}$$

we refer to this quantity as the *netput* process, although it does not correspond precisely to the netput processes constructed in the heavy traffic analysis of conventional queueing networks (e.g., Peterson 1991). Summing the inventory evolution equations (7) over all retailers and substituting the relevant definitions yields

$$Q(t) = \chi(t) - \frac{V}{\theta_T}I(t) + \epsilon_T(t). \tag{14}$$

The heavy traffic approximation for the IRP may be found as the limit (as $n \to \infty$) of a sequence of systems indexed by the heavy traffic parameter n. Even though no weak convergence proofs will be undertaken here, because some of the scalings that we introduce are non-traditional, we index quantities with n (in an appropriate place) to make the scalings clear. This indexing will be confined to this subsection; for the rest of the paper we leave off the index, with the understanding that we are considering a single system that has an associated value of n. The parameter n can be thought of as a large integer (e.g., 100) but (as is typically the case) the policy recommendations that emerge from our heavy traffic analysis are independent of n. The parameter n is used to normalize the various processes according to standard heavy traffic conventions (notice that only the process B_T undergoes a "fluid" scaling):

$$W_i^{(n)}(t) = \frac{Q_i^{(n)}(nt)}{\sqrt{n}} \quad \text{for all } i, \qquad W^{(n)}(t) = \sum_i W_i^{(n)}(t) = \frac{Q^{(n)}(nt)}{\sqrt{n}}, \tag{15}$$

$$Y^{(n)}(t) = \frac{I^{(n)}(nt)}{\sqrt{n}}, \qquad \hat{\chi}^{(n)}(t) = \frac{\chi^{(n)}(nt)}{\sqrt{n}}, \qquad \hat{\epsilon}_T^{(n)}(t) = \frac{\epsilon_T^{(n)}(nt)}{\sqrt{n}}, \tag{16}$$

$$\hat{\mathcal{D}}^{(n)}(t) = \frac{\mathcal{D}^{(n)}(nt)}{\sqrt{n}}, \qquad \hat{\mathcal{S}}_T^{(n)}(t) = \frac{\mathcal{S}_T^{(n)}(nt)}{\sqrt{n}} \quad \text{and} \quad \hat{B}_T^{(n)}(t) = \frac{B_T^{(n)}(nt)}{n}. \tag{17}$$

The processes $W^{(n)}$ and $Y^{(n)}$ represent the normalized inventory and idleness, respectively; to reduce the amount of notation, the normalized versions of the remaining processes contain

a "hat". Employing these scalings in (13) and (14) yields expressions for the normalized netput process

$$\hat{X}^{(n)}(t) = \sqrt{n}\left(\frac{V^{(n)}}{\theta_T^{(n)}} - \lambda^{(n)}\right)t + V^{(n)}\hat{\mathcal{S}}_T^{(n)}(\hat{B}_T^{(n)}(t)) - \hat{\mathcal{D}}^{(n)}(t), \tag{18}$$

and the normalized inventory process

$$W^{(n)}(t) = \hat{X}^{(n)}(t) - \frac{V^{(n)}}{\theta_T^{(n)}}Y^{(n)}(t) + \hat{\epsilon}_T^{(n)}(t). \tag{19}$$

To obtain a nontrivial control problem in heavy traffic, we normalize the system parameters in a particular fashion. The demand and inventory cost parameters are not scaled, and the other parameters are normalized as follows:

$$\hat{V}^{(n)} = \frac{V^{(n)}}{\sqrt{n}}, \tag{20}$$

$$\hat{\theta}_T^{(n)} = \frac{\theta_T^{(n)}}{\sqrt{n}}, \tag{21}$$

$$\mu_T^{(n)} = \sqrt{n}\left(\frac{V^{(n)}}{\theta_T^{(n)}} - \lambda^{(n)}\right) > 0, \tag{22}$$

$$\hat{c}_T^2(n) = \sqrt{n}c_T^2(n), \tag{23}$$

$$\hat{r}^{(n)} = \frac{r^{(n)}}{n}. \tag{24}$$

We assume that all the quantities on the left side of definitions (20)-(24) converge to finite and positive limits as $n \to \infty$. Equations (20)-(24) are the *heavy traffic conditions* and they specify, in a unified manner via the heavy traffic parameter n, the relative magnitudes of the various system parameters. These conditions are more extensive than those enforced in traditional queueing systems, and therefore warrant some discussion. Since the natural definition of the traffic intensity is $\rho_T = \lambda\theta_T/V$, condition (22) is the "traditional" heavy traffic condition, which requires that ρ_T be close to, but less than, unity.

Now we turn to conditions (20)-(21). Because the state space is compressed by a factor of $\sqrt{n}$ in the heavy traffic normalization, the vehicle capacity in terms of scaled inventory units is $V^{(n)}/\sqrt{n}$. Hence, if $V^{(n)}$ was $O(1)$ it would vanish in the limit, and our system would reduce to a variant of the multiclass make-to-stock queue analyzed in Wein (1992). Although such a model would be tractable, a limit that employs infinitesimal vehicle sizes

fails to capture the essence of the behavior of the original system. Therefore, we enforce condition (20), so that $V^{(n)}$ is $O(\sqrt{n})$ and the bulkiness of the retailer deliveries is retained in the limit. However, since the demand rate $\lambda^{(n)}$ is unscaled, we need to also scale the tour lengths according to (21) to ensure that the ratio $V^{(n)}/\theta_T^{(n)}$ converges to a finite and positive limit.

Turning to (23), note that since the vehicle capacity is $O(\sqrt{n})$, if the $c_T^2(n)$ is not scaled then a standard calculation shows that the variance term for the normalized netput process $\hat{X}^{(n)}$ is $O(n)$ and hence approaches infinity in the heavy traffic limit. Since $s_T^2(n) = c_T^2(n)\theta_T^2(n)$, by (21) and (23) we obtain $s_T^2(n) = \sqrt{n}\hat{s}_T^2(n)$, where $\hat{s}_T^2(n) = \hat{\theta}_T^2\hat{c}_T^2(n)$. Thus, by enforcing condition (23), we assume that the variance of the tour completion time is $O(\sqrt{n})$; in contrast, this quantity would be $O(n)$ if travel times were simply multiplied by $\sqrt{n}$. One way to achieve (23) is to assume that the travel time of the tour is the sum of $\sqrt{n}$ iid finite variance travel times. This construction could arise by superimposing the warehouse and retailer locations on a two-dimensional map with a fine grid, in such a way that the tour passes through approximately $\sqrt{n}$ grid points. However, this modeling artifice is problematic (because adjacent travel times would not likely be independent and the necessary data would be tedious to collect) and is not pursued here; see Rubio (1995) for further details.

Finally, we need to normalize the cost parameters to account for distortions in the relative magnitudes of the transportation and inventory costs that result from the heavy traffic scaling. The appropriate scaling is to allow the travel cost rate $r^{(n)}$ to be approximately n times larger than the inventory cost rates, as in condition (24); see Rubio for a detailed explanation.

In summary, the heavy traffic conditions assume that the vehicle must be busy the great majority of the time to meet average demand, the vehicle capacity must be large, the tour completion time must be large and nearly deterministic, and the travel cost rate must be very large relative to the inventory cost rates. The computational study in §6 reveals that our results are rather insensitive to these conditions.

2.3 System Behavior in Heavy Traffic

This subsection considers the limiting behavior of equations (18)-(19). Following Harrison (1988), we replace $\hat{B}_T(t)$, which is the fluid scaled busy time process, by $\rho_T t$; the justification for this substitution is that any policy that does not utilize the vehicle for a fraction ρ_T of the time over a sufficiently long time interval will generate extremely large inventory costs. In addition, we consider the normalized netput process embedded at tour completion epochs. Without some embedding or averaging the limit of the normalized netput process would not exist because it varies (after normalization) by $O(1)$ on a time of length $O(1/\sqrt{n})$. The process we consider is thus defined as

$$\tilde{\chi}(t) = \sqrt{n}\left(\frac{V}{\theta_T} - \lambda\right)\tau_{k-1} + V\hat{\mathcal{S}}_T(\rho_T\tau_{k-1}) - \hat{\mathcal{D}}(\tau_{k-1}) \text{ for } t \in [\tau_{k-1}, \tau_k) \ .$$

With this definition the standard tools of weak convergence (the functional central limit theorem for renewal processes, the random time change theorem and the continuous mapping theorem; see Billingsley 1968) can be used to show that the normalized netput process embedded at tour completion epochs $\tilde{\chi}$ is well approximated by a Brownian motion X with drift μ_T and variance $\sigma_T^2 = \lambda(c_d^2 + Vc_T^2)$.

Now we turn our attention to the process $\hat{\epsilon}_T$. This process equals zero at tour completion epochs and has jumps of size $O(1)$ whenever a delivery occurs and at the end of the cycle. In addition, because the tour length is $O(\sqrt{n})$ by (21), a tour takes only $O(1/\sqrt{n})$ time units under the heavy traffic normalization (where time is compressed by the factor n); hence, *tours occur instantaneously in the heavy traffic limit.* Consequently, neither $\hat{\epsilon}_T(t)$ nor the $m-$dimensional normalized inventory process converge to a limit in the usual sense. However, if we start with the heavy traffic normalization and expand time by a factor of $\sqrt{n}$, then a *fluid* scaling is obtained, where both time and space are compressed by the factor $\sqrt{n}$. At this faster time scale, the Brownian motion X remains constant, and the individual inventory levels move in a deterministic fashion, decreasing at a finite rate between the jumps at delivery epochs. The process $\hat{\epsilon}_T$ traverses through many tours before X changes value, and equals zero at each tour completion epoch.

This is similar to the state of affairs in the heavy traffic results of Coffman, Puhalskii

and Reiman. In their exhaustive polling system, the total queue length process behaves as a one-dimensional diffusion under the slow time scale associated with the heavy traffic scaling, and the individual queues move as a fluid under the faster time scale associated with the fluid limit. This *time scale decomposition* gives rise to a *heavy traffic averaging principle* (HTAP) that implies the following: for purposes of calculating performance measures for the individual queues, one can analyze the deterministic fluid cycle for each fixed value of the diffusion process. There are four key differences between the HTAP in the polling system and in the IRP. First, the fluid trajectories are different. In the polling problem, the fluid paths look like those for the economic production quantity (EPQ) model: they go up and down at a finite rate. In the IRP, the paths look like those from the economic order quantity (EOQ) model: they go down at finite rate but go up in jumps at delivery epochs. The second key difference relates to the issue of "control". In the polling system the exhaustive discipline guarantees that whenever the server switches from a queue, that queue is empty. This exerts a type of control that keeps the multidimensional process well behaved. There is no such natural mechanism in the IRP; we must introduce a dynamic allocation scheme to keep the multidimensional process well behaved. (This is done below.) Third, the time scale decomposition in the polling system emerged as a consequence of the standard heavy traffic normalization, whereas in the IRP it follows from the scaling assumptions (20)–(23). This gives rise to the fourth key difference — the proof of the HTAP. In the polling context a difficult proof involving a threshold queue was needed. For the IRP the proof follows from assumptions (20)–(23) and the properties of the dynamic control (we do not, however, provide the proof).

2.4 The Limiting Control Problem

As described above, the analysis of our limiting control problem decomposes onto two time scales. On the slow time scale associated with the diffusion limit, we can average out the effects of the controlled allocation process $\hat{c}_T$ and choose the vehicle idling policy. This policy is generated by the normalized cumulative idleness Y, which we assume is nondecreasing and

right continuous. Let $Z(t) = X(t) - \frac{V}{\theta_T}Y(t)$; this is the process that would be obtained if one were to observe the total inventory only at tour completion epochs. We refer to this process as the total *embedded* inventory process, to differentiate it from W.

At the faster time scale where the total embedded inventory is fixed at $Z(t) = x$ and the individual inventories behave as a fluid, we must find the optimal allocation policy that minimizes inventory costs per unit time. The limit cycle associated with an allocation policy can be viewed as a closed $m-$dimensional path, and the optimal allocation policy reduces to the problem of optimally placing a deterministic cycle in $\mathbf{R^m}$. Let $g(x)$ represent the inventory cost per unit time that is achieved by optimally locating a cycle when $Z(t) = x$.

We can now state the limiting stochastic control problem for the IRP with TSP routing: (i) find the optimal cycle placement for a given total embedded inventory level $Z(t) = x$, and its corresponding inventory cost rate $g(x)$; and (ii) choose the nondecreasing right continuous process Y to minimize

$$\limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T g(Z(t))dt - \hat{r}Y(T)\right] \tag{25}$$

$$\text{subject to} \quad Z(t) = X(t) - \frac{V}{\theta_T}Y(t). \tag{26}$$

The cycle placement problem is a nonlinear program and problem (25)-(26) is a singular control problem for Brownian motion; these two problems are solved in the next two subsections.

2.5 Optimal Cycle Placement and Dynamic Allocation

To optimally place the limit cycle, we follow the approach used in Markowitz, Reiman and Wein (1995) for the stochastic economic lot scheduling problem (ELSP). Let us fix $Z(t) = x$, and denote the individual fluid inventory levels by $\bar{W}_i(t) = Q_i(\sqrt{n}t)/\sqrt{n}$ (a "bar" will be used to denote quantities introduced for the fluid limit). The cycle placement can be defined in many ways and we choose to specify it by the vector $(x_1, x_2, \ldots, x_m)$, where x_i represents the lowest point during the cycle of $\bar{W}_i(t)$.

The choice of optimal $(x_1, \ldots, x_m)$ is a constrained optimization problem: we want to choose $(x_1, \ldots, x_m)$ to minimize the inventory cost rate subject to consistency with the total

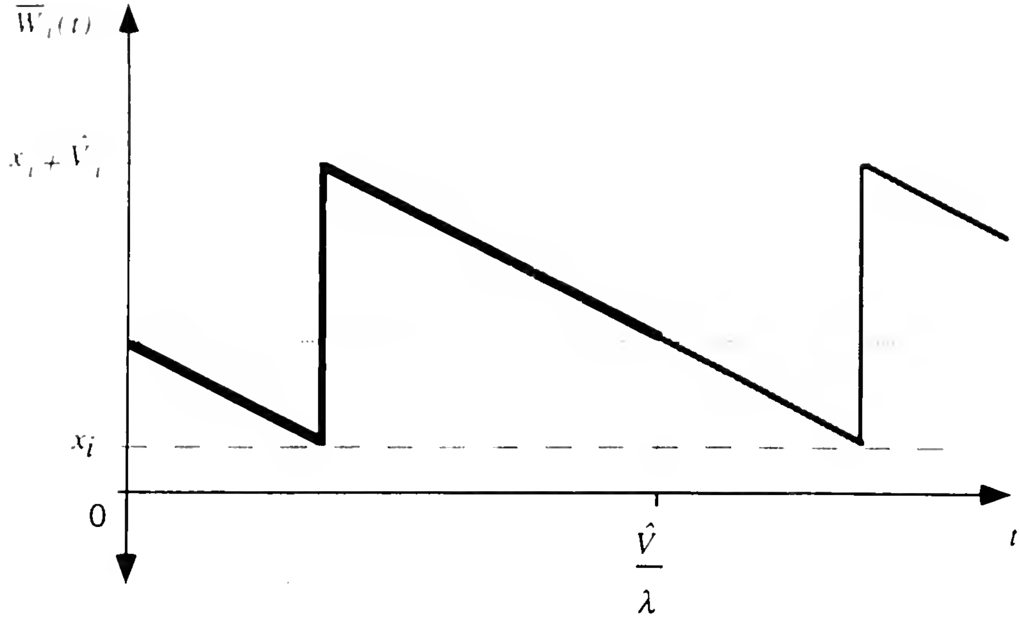

Figure 1: Fluid inventory evolution at retailer i during a nominal allocation cycle.

embedded inventory level. The inventory cost rate will come from the averaging principle to be described below. We first deal with the consistency issue.

To establish the relationship between the cycle placement variables x_i and the total embedded inventory level x, we need to introduce some new notation. Denote the mean travel time *along the TSP path* between any two sites $i, j \in 0, 1, \ldots, m$ by θ_{ij}^{TSP}; in terms of the inter-site mean travel times θ_{ij}, these quantities are defined by $\theta_{ij}^{TSP} = \sum_{k=i}^{j-1} \theta_{k,k+1}$ for $j > i$ and $\theta_{ii}^{TSP} = 0$. Since time is compressed by a factor of $\sqrt{n}$ in the fluid limit, define the corresponding travel times for the fluid model by $\bar{\theta}_{ij}^{TSP} = \theta_{ij}^{TSP}/\sqrt{n}$. If we measure time over a cycle so that the vehicle leaves the warehouse at $t = 0$, then $\bar{W}_i(0)$ is related to its corresponding cycle placement value x_i by (see Figure 1) $\bar{W}_i(0) = x_i + \lambda_i \bar{\theta}_{0i}^{TSP}$ for $i = 1, \ldots, m$. Summing these inventory levels over all retailers, we obtain

$$\sum_i x_i = x - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP}. \tag{27}$$

Given a vector $(x_1, \ldots, x_m)$ satisfying (27), we want to determine the associated inventory cost rate. The basic intuition of the time scale decomposition appears to provide the solution. Long term stability requires that, in the long run, the average amount delivered to retailer i per cycle be $V_i = \lambda_i V/\lambda$; under the diffusion and fluid scalings, this delivery size is given by $\hat{V}_i = V_i/\sqrt{n}$. Viewing the fluid inventory of retailer i in isolation with a

delivery of $\hat{V}_i$ on each tour cycle, we see a fluid starting out at $x_i + \hat{V}_i$ immediately after delivery, decreasing at a constant rate until x_i is reached just prior to the next delivery (see Figure 1). The inventory cost component associated with retailer i can then be calculated by considering the normalized inventory process W_i to be uniformly distributed on the interval $[x_i, x_i+\hat{V}_i]$. Although this approach provides the correct inventory cost rate, it turns out that a dynamic (state-dependent) delivery size is needed to keep the long run average cost finite. Simply delivering $\hat{V}_i$ to retailer i on every visit will result in an infinite long run average cost because this allocation leads to a null recurrent process. To see this, note that under this simple allocation scheme, the drift of $W_i(t)$ does not depend on $W_i(t)$ and equals zero, since the inventory "arrival" rate $\hat{V}_i/\hat{\theta}_T$ equals the demand rate λ_i when $\rho_T = 1$; with $\rho_T < 1$ a similar result is generated with the effective arrival rate being $\rho_T\hat{V}_i/\hat{\theta}_T$. With a zero drift, central limit theorem arguments indicate that the inventory or backlog will grow as $\sqrt{t}$. In fact, similar arguments can be used to explain some numerical results in Federgruen and Katalan (1994) and Wein, where a state-independent policy performs poorly in a stochastic setting.

A simple dynamic allocation policy escapes this difficulty. We determine delivery sizes at the warehouse as follows. Given a fluid inventory level $(w_1, \ldots, w_m)$ when the vehicle is at the warehouse, the fluid limit of the inventory immediately before delivery is $w_i - \lambda_i\bar{\theta}_{0i}^{TSP}$. If possible, we would like to deliver $d_i = x_i + \hat{V}_i - w_i + \lambda_i\bar{\theta}_{0i}^{TSP}$ to retailer i, in order to bring the fluid inventory level immediately after delivery to $x_i + \hat{V}_i$. If $d_i \geq 0$ for $i = 1, \ldots, m$ then this delivery allocation is feasible. If $d_i < 0$ for some i, then, since $\sum_i d_i = \hat{V}$, we must have $\sum_{\{i:d_i>0\}} d_i > \hat{V}$. This is a transient state for the fluid limit; within a finite number of cycles we will have $d_i \geq 0$ for all i. This transient interval has no effect on the long run average inventory cost, and an averaging principle will hold under this dynamic allocation scheme. In summary, the essence of the averaging principle here is that, under this dynamic allocation scheme, when $Z(t) = x$, $W_i(t)$ can be treated as if it is uniformly distributed between x_i and $x_i + \hat{V}_i$.

The average inventory cost per unit time is equal to the cost incurred over a cycle divided by the corresponding cycle length. The cost at retailer i may be obtained by simple

geometric arguments for any cycle placement x_i. When the cycle placement is sufficiently high (low) so that the inventory remains positive (negative) for the duration of the cycle, the cost is simply the holding (backordering) rate multiplied by the absolute value of $x_i + \hat{V}_i/2$, which is the average inventory level over a cycle. When the inventory changes sign during the cycle the total holding (backordering) cost over a cycle equals the area of one of the triangles above (below) the time axis multiplied by h_i (b_i). To obtain the time average inventory cost when there is a sign change we sum the areas of these two triangles and divide by the cycle length. In the heavy traffic limit, the amount of fluid delivered per cycle, $\hat{V}$, equals the amount demanded per cycle, which is $\lambda\hat{\theta}_T$; hence, we set the cycle length in the fluid model equal to $\hat{V}/\lambda$, rather than $\hat{\theta}_T$. In summary, we have the following expression for retailer i:

$$g_i(x_i) = \begin{cases} h_i(x_i + \frac{\hat{V}_i}{2}) & \text{if } x_i \geq 0 \\ \frac{h_i+b_i}{2\hat{V}_i}x_i^2 + h_i x_i + \frac{h_i\hat{V}_i}{2} & \text{if } -\hat{V}_i < x_i < 0 \\ -b_i(x_i + \frac{\hat{V}_i}{2}) & \text{if } x_i \leq -\hat{V}_i \end{cases} . \tag{28}$$

Notice that $g_i(x_i)$ is a convex function of x_i. With equation (28) in hand, the cycle placement problem is to minimize $\sum_i g_i(x_i)$ subject to (27).

Let us make the innocuous assumption that $b_i \geq h_i$ for all i, and define the labeling conventions $h_\ell = h = \min_i h_i$ and $b_p = b = \min_i b_i$, where $\ell = p$ is allowed. A closed-form solution to the cycle placement problem is found by using constraint (27) to turn the problem into one of unconstrained optimization over $m-1$ variables; readers are referred to an analogous optimization in Markowitz, Reiman and Wein for further details. The solution yields the vector of optimal placements x_i^* and $g(x)$, the inventory cost as a function of the total embedded inventory x. Not surprisingly, $g(x)$ is quadratic with linear edges in the inventory level x.

Proposition 1. *The solution to the cycle placement problem is*

$$\textit{Region 1:} \qquad x < \hat{\alpha}_T = \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} - \sum_i \frac{b+h_i}{b_i+h_i}\hat{V}_i,$$

$$x_i^* = -\frac{b+h_i}{b_i+h_i}\hat{V}_i \quad \textit{for } i \neq p.$$

$$
\begin{aligned}
x_p^* &= x - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} + \sum_{i \neq p} \frac{b + h_i}{b_i + h_i} \hat{V}_i \\
g(x) &= -bx + \hat{a}_1, \\
\hat{a}_1 &= b \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} + \frac{1}{2} \sum_i h_i \hat{V}_i - \frac{1}{2} \sum_i \frac{(b + h_i)^2}{b_i + h_i} \hat{V}_i; \\
\textit{Region 2:} \quad & \hat{\alpha}_T \leq x \leq \hat{\beta}_T = \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} - \sum_i \frac{h_i - h}{b_i + h_i} \hat{V}_i, \\
x_i^* &= \frac{2\hat{a}_2 \hat{V}_i}{h_i + b_i} \left(x - \sum_k \lambda_k \bar{\theta}_{0k}^{TSP} - \sum_k \frac{(h_i - h_k)\hat{V}_k}{b_k + h_k} \right), \\
g(x) &= \hat{a}_2 x^2 + \hat{a}_3 x + \hat{a}_4, \\
\hat{a}_2 &= \frac{1}{2} \left(\sum_i \frac{\hat{V}_i}{b_i + h_i} \right)^{-1}, \\
\hat{a}_3 &= 2\hat{a}_2 \left(\sum_i \frac{h_i \hat{V}_i}{b_i + h_i} - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} \right), \\
\hat{a}_4 &= \hat{a}_2 \left(\sum_i \frac{h_i \hat{V}_i}{b_i + h_i} - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} \right)^2 + \frac{1}{2} \sum_i \frac{b_i h_i \hat{V}_i}{b_i + h_i}; \\
\textit{Region 3:} \quad & x > \hat{\beta}_T, \\
x_i^* &= -\frac{h_i - h}{b_i + h_i} \hat{V}_i \ \textit{for } i \neq \ell, \\
x_\ell^* &= x - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} + \sum_i \frac{h_i - h}{b_i + h_i} \hat{V}_i, \\
g(x) &= hx + \hat{a}_5, \\
\hat{a}_5 &= -h \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} + \frac{1}{2} \sum_i h_i \hat{V}_i - \frac{1}{2} \sum_i \frac{(h_i - h)^2}{b_i + h_i} \hat{V}_i.
\end{aligned}
$$

In the region where the total inventory is much greater (smaller) than zero, the optimal cycle holds (backorders) most of the inventory at retailer ℓ (p), where it is cheapest to do so, while the cycle for the rest of the retailers remains close to zero. The exact level for each site depends upon two factors: the difference between its holding (backordering) cost h (b), and its nominal delivery size (or equivalently, the proportion of demand that the particular retailer represents). In the region where the total inventory is close to zero, the cycle placement at each retailer varies linearly with the embedded inventory in the system.

It is worth noting that, for the symmetric cost case (i.e., when $h_i = h$ and $b_i = b$ for

all $i = 1, \ldots, m$), the more natural solution of

$$x_i^* = \begin{cases} \frac{x - \sum_k \lambda_k \theta_{0k}^{TSP}}{m} & \text{if} \quad x \notin [\hat{\alpha}_T, \hat{\beta}_T] \\ \frac{\lambda_i}{\lambda} \left(x - \sum_k \lambda_k \theta_{0k}^{TSP} \right) & \text{if} \quad \hat{\alpha}_T \le x \le \hat{\beta}_T \end{cases}$$

is also optimal.

2.6 Optimal Base Stock Level

Now that $g(x)$ is known, we can proceed with the solution to the one-dimensional Brownian control problem. The following proposition is proved in Appendix A of Rubio.

Proposition 2. *The optimal solution to (25)-(26) is* $Y^*(t) = \sup_{0 \le s \le t}\{X(s) - z_T^*\}^+$ *for some base stock level* z_T^*.

Hence, the optimal solution is the local time of the Brownian motion at the barrier z_T^*, and the optimally controlled process Z is a reflected Brownian motion (RBM) on $(-\infty, z_T^*]$ (see §2.2 of Harrison 1985 for a definition). The remainder of this subsection is devoted to the derivation of z_T^*, which can be found by using two well known facts regarding an RBM on $(-\infty, z]$. First, for Y defined in Proposition 2 we have $\lim_{t \to \infty} t^{-1} E_x[Y(t)] = \theta_T \mu_T / V$ for $\mu_T > 0$, which is independent of the base stock level z. Hence, the transportation cost does not affect the selection of z, and the problem simplifies to minimizing $\limsup_{T \to \infty} \frac{1}{T} E\left[\int_0^T g(Z(t))dt\right]$ subject to (26).

The steady state density for Z is given by $p_Z(x) = \hat{\nu}_T e^{\hat{\nu}_T(x-z)}$ if $x \le z$ and $p_Z(x) = 0$ if $x > z$, where $\hat{\nu}_T = 2\mu_T/\sigma_T^2 > 0$. Therefore, the optimal base stock level can be found by minimizing

$$\begin{aligned} \hat{F}_T(z) &= \int_{-\infty}^{\hat{\alpha}_T} (-bx + \hat{a}_1)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx + \int_{\hat{\alpha}_T}^{\hat{\beta}_T} (\hat{a}_2 x^2 + \hat{a}_3 x + \hat{a}_4)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx \\ &\quad + \int_{\hat{\beta}_T}^{z} (hx + \hat{a}_5)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx \qquad \text{for} \quad z \ge \hat{\beta}_T, \quad \text{and} \end{aligned} \tag{29}$$

$$\hat{F}_T(z) = \int_{-\infty}^{\hat{\alpha}_T} (-bx + \hat{a}_1)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx + \int_{\hat{\alpha}_T}^{z} (\hat{a}_2 x^2 + \hat{a}_3 x + \hat{a}_4)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx \tag{30}$$

for $\hat{\alpha}_T < z < \hat{\beta}_T$. The constants in (29) and (30) have the same definitions as in §2.5. Note that while the optimal base stock level z_T^* always satisfies $z_T^* > \hat{\alpha}_T$ (this is easily seen by the fact that $g(x)$ is linear and has a negative slope for $x < \hat{\alpha}_T$), it need not be larger than $\hat{\beta}_T$.

The following proposition is derived by using integration by parts on (29) and (30), and then taking the first two derivatives of $\hat{F}_T(z)$ with respect to z.

Proposition 3. *The value that minimizes $\hat{F}_T(z)$ is*

$$z_T^* = -\frac{1}{\hat{\nu}_T} \ln \left[\left(\frac{h}{b+h} \right) \left(\frac{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)}{e^{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)} - 1} \right) \right] + \hat{\alpha}_T \quad \textit{if} \quad z_T^* \geq \hat{\beta}_T; \tag{31}$$

otherwise, z_T^* is the solution to

$$\frac{2\hat{a}_2}{\hat{\nu}_T} e^{-\hat{\nu}_T(z_T - \hat{\alpha}_T)} + 2\hat{a}_2 z_T + \hat{a}_3 - \frac{2\hat{a}_2}{\hat{\nu}_T} = 0. \tag{32}$$

Furthermore, the predicted optimal cost is $\hat{F}_T(z_T^*) = hz_T^* + \hat{a}_5$ if $z_T^* \geq \hat{\beta}_T$, and $\hat{F}_T(z_T^*) = \hat{a}_2(z_T^*)^2 + \hat{a}_3 z_T^* + \hat{a}_4$ otherwise.

One can show (from the fact that $\hat{F}_T(z)$ is convex and continuously differentiable) that there is a unique optimum base stock level; that is, either there exists a solution z_T^* to (31) that satisfies $z_T^* \geq \hat{\beta}_T$ or a solution z_T^* to (32) that satisfies $z_T^* < \hat{\beta}_T$, but not both.

2.7 The Proposed Policy

In this subsection we map the solution of the approximating heavy traffic control problem into a policy for the original IRP with TSP routing. The control concerns two decisions: whether the vehicle should be busy or idle, and how to assign the load among the retailers during a tour. We address the load allocations first.

Since the system evolves dynamically in time, the decision of how much of the load to leave at each retailer is best delayed until the vehicle arrives at the site. Let t_0 correspond to the epoch at which the vehicle leaves the warehouse with a full load, and consider the epoch $t_i^- > t_0$, which is the point in time just before the vehicle arrives at retailer i. At time t_i^-, the state of the system is given by the inventory levels at the retailers, $\big(Q_1(t_i^-), \ldots, Q_m(t_i^-)\big)$, and the size of the remaining load, $L(t_i^-)$. The mapping from heavy traffic solution to proposed policy is straightforward: the proposed policy attempts to track the heavy traffic solution (in particular, the optimal cycle placement) as closely as possible. The key issue to be addressed is that the heavy traffic solution is expressed in terms of normalized space

and time and in terms of the total embedded inventory process, whereas the proposed policy must be expressed in terms of $(Q_1(t_i^-), \ldots, Q_m(t_i^-), L(t_i^-))$.

Recall that equation (27) relates the scaled cycle placement vector x_i and the normalized total embedded inventory $Z(t) = x$. Since the load allocation decision is taken when each retailer is reached, we first establish a relationship between the current total system inventory $Q(t_i^-) = \sum_j Q_j(t_i^-)$ and the corresponding embedded inventory level. Because we need to reverse the scalings in the solution to the heavy traffic control problem, let us define the unscaled embedded inventory $q = \sqrt{n}x$ and the unscaled cycle placement vector $q_i = \sqrt{n}x_i$. In keeping with the behavior predicted by the heavy traffic averaging principle, we develop this relation under a deterministic evolution for the retailer inventories over the course of a cycle. If the vehicle leaves the warehouse at time t_0, then it arrives at retailer i at time t_i^-, where $t_i = t_0 + \theta_{0i}^{TSP}$. Therefore, the retailer inventories relate to the cycle placement parameters by $Q_j(t_i^-) = q_j + \lambda_j \theta_{ij}^{TSP}$ for $j \geq i$ and $Q_j(t_i^-) = q_j + V_j - \lambda_j \theta_{ji}^{TSP}$ for $j < i$. Summing over all retailers, we get

$$Q(t_i^-) = \eta_i + \sum_j q_j, \tag{33}$$

where $\eta_i = \sum_{j<i}(V_j - \lambda_j \theta_{ji}^{TSP}) + \sum_{j \geq i} \lambda_j \theta_{ij}^{TSP}$ is an epoch locator constant for retailer i. Making the substitutions $q_i/\sqrt{n} = x_i$, $q/\sqrt{n} = x$ and $\theta_{ij}^{TSP}/\sqrt{n} = \bar{\theta}_{ij}^{TSP}$ into (27) yields the unscaled version of constraint (27),

$$\sum_i q_i = q - \sum_i \lambda_i \theta_{0i}^{TSP}. \tag{34}$$

Using equations (33) and (34), we can express the total inventory at time t_0 (i.e., when the vehicle was at the warehouse) as a translation of the inventory vector at time t_i^-:

$$Q(t_0) = q = \sum_j Q_j(t_i^-) - \eta_i + \sum_j \lambda_j \theta_{0j}^{TSP} \quad \text{for } i = 0, \ldots, m. \tag{35}$$

This equation maps the current inventory levels $Q_j(t_i^-)$ into the one-dimensional quantity q that is required to interpret the heavy traffic results. In particular, for a given value of q we can find $q_i^* = \sqrt{n}x_i^*$, the corresponding optimal (unscaled) cycle placement parameters by reversing the normalizations in Proposition 1. As before, this is done by letting $q_i/\sqrt{n} = x_i$,

$q/\sqrt{n} = x$, $V_i/\sqrt{n} = \hat{V}_i$, and $\theta_{ij}^{TSP}/\sqrt{n} = \bar{\theta}_{ij}^{TSP}$ to obtain the following expressions for q_i^* in terms of the state q:

$$\text{Region 1:} \qquad q < \alpha_T = \sum_i \lambda_i \theta_{0i}^{TSP} - \sum_i \frac{b+h_i}{b_i+h_i} V_i, \tag{36}$$

$$q_i^* = -\frac{b+h_i}{b_i+h_i} V_i \text{ for } i \neq p, \tag{37}$$

$$q_p^* = q - \sum_i \lambda_i \theta_{0i}^{TSP} + \sum_{i \neq p} \frac{b+h_i}{b_i+h_i} V_i; \tag{38}$$

$$\text{Region 2:} \qquad \alpha_T \le q \le \beta_T = \sum_i \lambda_i \theta_{0i}^{TSP} - \sum_i \frac{h_i - h}{b_i+h_i} V_i, \tag{39}$$

$$q_i^* = \frac{2a_2 V_i}{h_i + b_i}\left(q - \sum_k \lambda_k \theta_{0k}^{TSP} - \sum_k \frac{(h_i - h_k)V_k}{b_k + h_k}\right), \tag{40}$$

$$a_2 = \frac{1}{2}\left(\sum_i \frac{V_i}{b_i + h_i}\right)^{-1}; \tag{41}$$

$$\text{Region 3:} \qquad \beta_T < q, \tag{42}$$

$$q_i^* = -\frac{h_i - h}{b_i + h_i} V_i \text{ for } i \neq \ell, \tag{43}$$

$$q_\ell^* = q - \sum_i \lambda_i \theta_{0i}^{TSP} + \sum_i \frac{h_i - h}{b_i + h_i} V_i, \tag{44}$$

which are independent of the scaling parameter n.

We can now use these results to determine the delivery size at retailer i. Under the deterministic inventory evolution for the optimal cycle placement, $Q_i(t_i^+)$ (i.e. the inventory level at retailer i just after the delivery is made) satisfies

$$Q_i(t_i^+) = q_i^* + V_i. \tag{45}$$

If we deliver d_i units to retailer i then the actual inventory after the delivery is made is $Q_i(t_i^-) + d_i$; equating this quantity to (45) yields the desired delivery size,

$$d_i = q_i^* + V_i - Q_i(t_i^-). \tag{46}$$

Because we cannot allocate more than the available load and do not want to make negative deliveries, the proposed delivery size is given by

$$d_i^* = \max[d_i, 0] + \min[0, L(t_i^-) - d_i] \text{ for } i = 1, 2, \ldots, m-1. \tag{47}$$

Finally, to guarantee that the vehicle returns to the warehouse empty, we set

$$d_m^* = L(t_m^-). \tag{48}$$

We could in principle allow negative deliveries, as long as there is inventory available at retailer i and the total amount of load the vehicle carries as it leaves this retailer is kept under its total capacity. However, because the vehicle returns empty, the items accrued from negative deliveries would most likely be shifted to the last few retailers of the tour, which will not necessarily bring the state of the system closer to the optimal cycle; hence, we disallow negative deliveries.

To recapitulate, the proposed dynamic delivery allocations are derived by the following procedure: (i) observe the current inventory levels $(Q_1(t_i^-), \ldots, Q_m(t_i^-))$ and compute the unscaled embedded inventory q via (35), (ii) use (36)-(44) to derive the optimal unscaled cycle placement parameters q_i^*, and (iii) observe the current remaining load $L(t_i^-)$ and compute the proposed delivery size via (46)-(48).

We now turn our attention to the busy/idle policy, which has decision epochs when the vehicle is at the warehouse. At these points in time, the vehicle starts a new tour if the total inventory level $\sum_j Q_j(t)$ is below the unscaled aggregate base stock level $q_T^* = \sqrt{n}\, z_T^*$; otherwise, it idles. Reversing the normalizations in Proposition 3 yields the optimal unscaled base stock level solely in terms of the original problem parameters:

$$q_T^* = -\frac{1}{\nu_T} \ln\left[\left(\frac{h}{b+h}\right)\left(\frac{\nu_T(\beta_T - \alpha_T)}{e^{\nu_T(\beta_T - \alpha_T)} - 1}\right)\right] + \alpha_T \quad \text{if} \quad q_T^* \geq \beta_T; \tag{49}$$

otherwise, q_T^* is the solution to

$$\frac{2a_2}{\nu_T} e^{-\nu_T(q_T - \alpha_T)} + 2a_2 q_T + a_3 - \frac{2a_2}{\nu_T} = 0. \tag{50}$$

Furthermore, the predicted optimal cost is given by $F_T(q_T^*) = hq_T^* + a_5$ if $q_T^* \geq \beta_T$, and by $F_T(q_T^*) = a_2(q_T^*)^2 + a_3 q_T^* + a_4$ otherwise. The constants in these expressions are the unscaled counterparts of the ones defined earlier:

$$\nu_T = \frac{2(1-\rho_T)V}{\lambda\theta_T(c_d^2 + Vc_T^2)},$$

$$a_3 = 2a_2\left(\sum_i \frac{h_i V_i}{b_i + h_i} - \sum_i \lambda_i \theta_{0i}^{TSP}\right),$$

$$a_4 = a_2\left(\sum_i \frac{h_i V_i}{b_i + h_i} - \sum_i \lambda_i \theta_{0i}^{TSP}\right)^2 + \frac{1}{2}\sum_i \frac{b_i h_i V_i}{b_i + h_i} \quad \text{and}$$

$$a_5 = -h\sum_i \lambda_i \theta_{0i}^{TSP} + \frac{1}{2}\sum_i h_i V_i - \frac{1}{2}\sum_i \frac{(h_i - h)^2}{b_i + h_i} V_i.$$

We have completely characterized a dynamic control policy that depends exclusively on the original system parameters. The policy specifies the two controllable aspects of the system: the aggregate base stock level defined in (49)-(50) determines the vehicle idling policy, and the delivery sizes d_i^* defined in (47)-(48) characterize the allocation of units to retailers.

3 The IRP with Direct Shipping

The only difference between the direct shipping (DS) case and the TSP case is the routing scheme used when the vehicle is operating. We retain all notation from §2, occasionally using the subscript "D" (for direct shipping) in place of the subscript "T" (for TSP). Moreover, since the procedure is very similar in both cases, we omit nearly all of the details for the DS case, describing only the distinctive aspects of the analysis. The most significant difference between DS and TSP is that the DS case does not have a cyclic structure. This has several consequences, one of which is that the results are not as theoretically solid as in the TSP case.

3.1 Heavy Traffic Analysis

In the DS case, the vehicle always leaves the warehouse with a full load, visits a single retailer and returns empty, so that every time a retail site is visited its inventory level increases by V units. As before, it is convenient to express the dynamic allocation as deviations from a nominal policy. The nominal policy we consider is not achievable: we assume that under the nominal policy an amount V_i is delivered to retailer i in every delivery. We let $S_D(t)$ denote the number of DS deliveries made by a continuously active vehicle during $[0, t]$. We

then let $c_i^D(t)$ be defined by the analog of equation [illegible] obtained by replacing T's by D's. Then $c_D(t) = \sum_i c_i^D(t)$ is the total amount delivered during the current trip. These processes satisfy (2)–(6). Equation (7) still holds as well, once T's are replaced by D's. The problem formulation is thus nearly identical to (2)–(12).

The DS case lacks the natural cyclic structure of the TSP. Tour times in the TSP are i.i.d., and each tour results in the delivery of V units. The DS case would have a cyclic structure if the sequence of retailers visited followed a cyclic pattern (such as a polling table), or had a regenerative structure (such as a Markov chain). Neither of these can be used in the DS case, for exactly the same reason that fixed delivery sizes could not be used in the TSP case: inventory costs would be infinite over the long run. A dynamic policy, described in §3.2, is used. For this policy the fraction of total shipments that go to retailer i does not vary over times of order n. To satisfy average demand at all retailers, this fraction must be λ_i/λ for retailer i. The traffic intensity for the DS system is thus $\rho_D = 2\sum_i \lambda_i \theta_{0i}/V$.

We use the same heavy traffic normalizations as in §2.2. The heavy traffic conditions are given by (20)–(24), with (21) and (23) understood to hold, respectively, for θ_{0i} and c_{0i}^2, $1 \leq i \leq m$, and (22) replaced by

$$\mu_D = \sqrt{n}\left(\frac{\lambda V}{2\sum_i \lambda_i \theta_{0i}} - \lambda\right) > 0 \ . \tag{51}$$

As mentioned above, the DS case does not have the cyclic structure of the TSP case. Thus we cannot use functional central limit theorems based on renewal processes to prove convergence to a Brownian motion. In the DS case the parameter corresponding to s_T^2 in the TSP case, s_D^2, is given by $s_D^2 = \lambda^{-1}\sum_i \lambda_i \theta_{0i}^2 c_{0i}^2$. This expression would clearly arise if the next retailer were chosen in an i.i.d. manner. It can be shown to hold for any policy where the fraction of times that a retailer is visited does not vary over times of order n using the Random Time Change Theorem.

The lack of a cyclic structure makes it impossible to consider an embedded normalized netput process. Indeed, if we embed at epochs during which the vehicle is at the warehouse, we will not obtain a meaningful process. We must thus average the normalized netput process to obtain a meaningful limit. By arguments similar to those in §2.3, this averaged

normalized netput process is well approximated by a Brownian motion X with drift μ_D and variance

$$\sigma_D^2 = \lambda \left(c_d^2 + \frac{\lambda V \sum_i \lambda_i \theta_{0i}^2 c_{0i}^2}{2(\sum_i \lambda_i \theta_{0i})^2} \right) .$$

The averaged total inventory process is defined by

$$Z(t) = X(t) - \frac{\lambda V}{2 \sum_i \lambda_i \theta_{0i}} Y(t) . \tag{52}$$

Now we slow down time by a factor of $\sqrt{n}$ and turn to the fluid model. Again we face the problem of optimally placing the limit cycles for the deterministic evolution of the retailer inventories in $\mathbf{R^m}$. In contrast to the TSP case, we introduce an approximation to facilitate this optimization. We motivate this approximation by use of an example. Suppose that $V = 60$, $\lambda_1 = 3$, $\lambda_2 = 2$ and each retailer was exactly six time units away from the warehouse. Then a continually busy vehicle using the polling table (12121) could visit, *on average*, retailer 1 every 20 time units and retailer 2 every 30 time units, thereby satisfying average demand. Although optimally placing a limit cycle for a small polling table such as this one is manageable, the optimization problem gets unwieldy very quickly as the size of the table grows. Our approximation assumes the existence of an *idealized* policy that would make a delivery to retailer i every V/λ_i time units in the fluid model; in the context of our example, we assume that retailer 1 (2) receives a delivery *exactly* every 20 (30) time units, even though the (12121) polling table cannot achieve such perfect regularity in the fluid model.

Hence, our approach is to optimally place an idealized fluid cycle at the fast time scale, and then track this cycle as closely as possible with our proposed policy. Because deliveries are perfectly regular in the idealized cycle, the use of this approximation causes us to underestimate the inventory cost incurred over a cycle; however, simulation results in §6.1 show that the heavy traffic analysis incorporating this approximation appears to be very accurate, at least for the five-retailer cases considered there. Moreover, the use of an idealized cycle allows us to avoid the task of determining the actual behavior of the individual inventory levels on the fluid time scale. This task is more than just tedious: due to the lack of a cyclic structure it appears to be extremely difficult.

Now we turn to the optimal placement of the idealized cycle. We still define the cycle placement by the vector $(x_1, x_2, \ldots, x_m)$, where x_i represents the lowest point during the cycle of the fluid inventory level at retailer i. Under the idealized policy, the fluid inventories are similar to the TSP paths pictured in Figure 1: the only differences are the delivery size (in this case we deliver a full load of $\hat{V}$ units on each visit to a retailer) and the visit frequency, which equals $\hat{V}/\lambda_i$ in order to maintain a balanced flow.

The next step is to establish the relationship between the cycle placement variables x_i and the averaged total inventory level $Z(t) = x$. The constraint related to consistency between individual and total inventory levels takes the form that the averaged total inventory equals the sum of the average individual inventory levels. The average fluid inventory at retailer i over an idealized cycle is $x_i + \hat{V}/2$. Hence, when the averaged total inventory $Z(t) = x$, the cycle placement parameter must satisfy

$$\sum_i x_i = x - \frac{m\hat{V}}{2}. \tag{53}$$

Because the fluid delivery size equals $\hat{V}$ under the DS case, the inventory cost function $g_i(x_i)$ is given as in (28), except that $\hat{V}_i$ is replaced by $\hat{V}$. Comparing constraints (27) and (53), it is clear that the optimal cycle placement is precisely the solution given in the TSP case, except that we replace $\hat{V}_i$ by $\hat{V}$ and $\sum_i \lambda_i \bar{\theta}_{0i}^{TSP}$ by $m\hat{V}/2$.

The computation of the optimal vehicle idling policy is identical to that in §2, except that the parameter of the exponential stationary density of the RBM is $\hat{\nu}_D = 2\mu_D/\sigma_D^2$. Hence, Proposition 3 characterizes the optimal base stock level for the DS case, with $\hat{\nu}_D$ replacing $\hat{\nu}_T$, and with the substitutions described above in the corresponding definitions of the constants $\hat{a}_i$ and the thresholds $\hat{\alpha}$ and $\hat{\beta}$.

3.2 The Proposed Policy

The mapping from heavy traffic solution to proposed policy uses the same philosophy as in the TSP case. We begin by establishing a relationship between the total inventory in the system at the current decision epoch and the cycle placement parameters. Although there are several possible ways to do this, we keep track of the vector process $(r_1(t), \ldots, r_m(t))$,

which specifies the time of the most recent visit to each retailer. Hence, if we denote the current time by t, then $t - r_i(t)$ represents the elapsed time since the vehicle last visited retailer i. The inventory at retailer i at time t relates to the unscaled cycle placement parameter $q_i = \sqrt{n}\, x_i$ via

$$Q_i(t) = q_i + V - \lambda_i(t - r_i(t)). \tag{54}$$

Since stochastic effects can lead to unusually long intervisit periods, it is possible to have $V - \lambda_i(t - r_i(t)) < 0$, which would make the cycle placement of the retailer higher than the current inventory level, thereby contradicting the definition of q_i. Because one would expect that $Q_i(t) \geq q_i + \lambda_i\theta_{0i}$, we modify equation (54) to $Q_i(t) = q_i + \max[V - \lambda_i(t - r_i(t)), \lambda_i\theta_{0i}]$; this modification leads to considerable improvements in system performance in the simulation study. Summing over all retailers we obtain $\sum_i q_i = Q(t) - u(t)$, where $u(t) = \sum_i \max\left[V - \lambda_i(t - r_i(t)), \lambda_i\theta_{0i}\right]$. Combining this equation with the unscaled version of (53) yields

$$q = \sum_i Q_i(t) - u(t) + \frac{mV}{2}, \tag{55}$$

which relates the unscaled averaged inventory q (this quantity represents the unscaled average total inventory in the DS case) to the current inventory level. Reversing the heavy traffic normalizations, we obtain the following formulas for the unscaled optimal cycle placement vector $q_i^* = \sqrt{n}\, x_i^*$ given q:

$$\begin{aligned}
&\text{Region 1:} && q < \alpha_D = -V\sum_i \frac{b + h_i}{b_i + h_i} + \frac{mV}{2},\\
&q_i^* &=\ & -\frac{b + h_i}{b_i + h_i}V \quad \text{for } i \neq p,\\
&q_p^* &=\ & q - \frac{mV}{2} + V\sum_{i\neq p}\frac{b + h_i}{b_i + h_i};\\
&\text{Region 2:} && \alpha_D \leq q \leq \beta_D = V\sum_i \frac{h - h_i}{b_i + h_i} + \frac{mV}{2},\\
&q_i^* &=\ & \frac{2a_7V}{h_i + b_i}\left(q - \frac{mV}{2} - V\sum_k \frac{h_i - h_k}{b_k + h_k}\right),\\
&a_7 &=\ & \frac{1}{2V}\left(\sum_i \frac{1}{b_i + h_i}\right)^{-1};\\
&\text{Region 3:} && \beta_D < q,
\end{aligned}$$

$$
\begin{aligned}
q_i^* &= -\frac{h_i - h}{b_i + h_i} V \text{ for } i \neq \ell, \\
q_\ell^* &= x - \frac{mV}{2} + V \sum_i \frac{h_i - h}{b_i + h_i},
\end{aligned}
$$

which are independent of the scaling factor n.

We use this cycle placement as an "ideal" m-dimensional inventory state for a given q and $u(t)$, and choose the next retailer so as to bring the current inventory vector as close as possible to this ideal state. Let t_0 be the time epoch at which the vehicle is ready to depart from the warehouse, and consider the inventory evolution over the next delivery trip. Under a deterministic inventory evolution, the vehicle will reach retailer i (if it chooses to go there next) at time $t_i = t_0 + \theta_{0i}$. In the deterministic tour corresponding to the optimal cycle placement, the retailer inventory levels right after a delivery is made to retailer i is given by $Q_i^*(t_i^+) = q_i^* + V$ and

$$
Q_j^*(t_i^+) = \max\left[q_j^* + V - \lambda_j(t_0 + \theta_{0i} - r_j(t_0)),\ q_j^* + \lambda_j(\theta_{0i} + \theta_{0j})\right] \quad \text{for } j \neq i,
$$

where the maximization makes the adjustment for long intervisit times as discussed above. In contrast, under the deterministic evolution the actual inventory vector after a delivery to retailer i is given by $Q_i(t_i^+) = Q_i(t_0) + V - \lambda_i\theta_{0i}$ and $Q_j(t_i^+) = Q_j(t_0) - \lambda_j\theta_{0i}$ for $j \neq i$. Therefore, the resulting Euclidean distance between the ideal and actual inventory vectors after a delivery to retailer i is $\Delta(i) = \sqrt{\sum_j \left(Q_j(t_i^+) - Q_j^*(t_i^+)\right)^2}$. The proposed control sends the vehicle to retailer k, where $k = \arg\min_i \Delta(i)$.

Finally, as in the TSP case, the busy/idle control is a direct unscaling of the heavy traffic results. The only added complexity is that for the DS case the vehicle visits the warehouse after every delivery and so has many possible idling decision epochs. By equation (55), our proposed policy idles a vehicle at the warehouse whenever

$$
Q(t) - u(t) + \frac{mV}{2} > w_D^*,
$$

where $w_D^* = \sqrt{n}\, z_D^*$ is the unscaled idling threshold. Making the suitable scaling substitutions, we have that

$$
w_D^* = -\frac{1}{\nu_D} \ln\left[\left(\frac{h}{b+h}\right)\left(\frac{\nu_D(\beta_D - \alpha_D)}{e^{\nu_D(\beta_D - \alpha_D)} - 1}\right)\right] + \alpha_D \quad \text{if} \quad w_D^* \geq \beta_D; \tag{56}
$$

otherwise, w_D^* is the solution to

$$\frac{1}{\nu_D} e^{-\nu_D(w_D-\alpha_D)} + w_D - \frac{1}{\nu_D} - \frac{mV}{2} + V \sum_i \frac{h_i}{b_i + h_i} = 0. \tag{57}$$

Furthermore, the predicted optimal cost is given by

$$F_D(w_D^*) = hw_D^* - h\frac{mV}{2} + \frac{V}{2}\sum_i h_i - \frac{V}{2}\sum_i \frac{(h_i - h)^2}{b_i + h_i} \quad \text{if} \quad w_D^* \geq \beta_D,$$

and $F_D(w_D^*) = a_7(w_D^*)^2 + a_8 w_D^* + a_9$ otherwise. The constants in these equations are

$$\nu_D = \frac{(1-\rho_D)\lambda V}{\sigma_D^2 \sum_i \lambda_i \theta_{0i}},$$

$$a_8 = 2Va_7\left(\sum_i \frac{h_i}{b_i + h_i} - \frac{m}{2}\right) \quad \text{and}$$

$$a_9 = a_7\left(V\sum_i \frac{h_i}{b_i + h_i} - \frac{mV}{2}\right)^2 + \frac{V}{2}\sum_i \frac{b_i h_i}{b_i + h_i},$$

where (α_D, β_D, a_7) are defined in the unscaled cycle placement formulas above.

4 Comparison of TSP and DS Routing

In this section we compare the relative performance of the two fixed routing schemes (TSP and DS). The predicted cost functions F_T, F_D derived in §2 and §3 represent only the inventory component of the system cost. Denote a generic fixed routing scheme by $\Re \in \{\text{TSP}, \text{DS}\}$, and by $C(\Re)$ the total cost for the system under this scheme. The total system cost is obtained by adding the transportation cost (or equivalently, subtract the idleness reward) to the inventory cost; that is, we set $C(\Re) = F_\Re(w_\Re^*) - r(1-\rho_\Re)$. Notice that the transportation cost is independent of the stochastic nature of the system.

A crucial observation is that $\rho_D < \rho_T$; this fact (see Rubio for a proof) is a simple consequence of the triangle inequality. Although trivial to prove, this inequality has several important implications. First, the DS policy achieves lower transportation costs than the TSP policy. This is quite interesting since, at first glance, one might expect the converse to hold. However, minimization of the steady state transportation cost in the IRP context

is equivalent to maximizing the amount of items delivered per unit time traveled; hence, full load direct shipping provides the highest transportation efficiency of *any* fixed routing scheme. More importantly, for any given problem instance, the demand rate can be increased until $\rho_T = 1$ and $\rho_D < 1$; that is, there exist some demand levels where the DS policy would be stable while the TSP policy would not. We should note that $\rho_{\Re} < 1$ is a necessary condition for stability of any fixed routing scheme $\Re$ but it is not sufficient. In particular, having $\rho_{\Re} < 1$ will keep the total inventory stable but, in the absence of adequate dynamic load allocation, it is possible to accumulate inventory at one retail site while backorders grow without bound at another. Hence DS will dominate TSP routing as $\rho_T \to 1$ as long as some form of stable dynamic allocation is used in the DS case.

The remainder of this section investigates the relative performance of the DS and TSP schemes as a function of the cost parameters r and b. However, readers should keep in mind that while the qualitative statements below are true, these results are not exact, because our calculation of $F_D(w_D^*)$ is approximate and represents a slight underestimate of the true heavy traffic cost under the DS policy. The inequality $\rho_D < \rho_T$ implies that DS is preferred to the TSP policy if the transportation cost is high enough. In particular, the DS policy achieves a lower overall cost for any $r > (F_T(w_T^*) - F_D(w_D^*))/(\rho_T - \rho_D)$. While the value of this threshold cost may be found numerically for any particular problem instance, a more precise characterization requires a better understanding of the relationship between the inventory costs in both systems. Unfortunately, it is hard to make simple inventory cost performance comparisons for the different routing schemes, primarily because the base stock levels, and hence the predicted inventory costs, are not in closed form (see equations (50) and (57)).

To study the relative inventory cost performance of the TSP and DS schemes, let us consider the case where the inventory costs at the retailers are symmetric (i.e., $h_i = h$ and $b_i = b$ for all i) and b becomes large. Because the value of w_T^* and w_D^* in (49) and (56) is increasing in b/h, one expects that there exist some critical values b_T, b_D such that if b is increased above them (while leaving h fixed) the optimal base stock is given in closed form. These critical values indeed exist and, for the case of symmetric costs, have the following

closed form expressions:

$$b_T = h\left[\frac{\nu_T V e^{\nu_T V}}{e^{\nu_T V} - 1} - 1\right] \quad \text{and} \quad b_D = h\left[\frac{\nu_D m V e^{\nu_D m V}}{e^{\nu_D m V} - 1} - 1\right].$$

For $b > \max\{b_T, b_D\}$, the inventory cost difference $F_T(w_T^*) - F_D(w_D^*)$ can be expressed as

$$h\left(\frac{\nu_D - \nu_T}{\nu_D \nu_T}\ln\left[1 + \frac{b}{h}\right] + \frac{1}{\nu_T}\ln\left[\frac{e^{\nu_T V} - 1}{\nu_T V}\right] - \frac{1}{\nu_D}\ln\left[\frac{e^{\nu_D m V} - 1}{\nu_D m V}\right] + \frac{(m-1)V}{2}\right). \tag{58}$$

As $b \to \infty$ the value of (58) is dominated by the term $(\nu_T^{-1} - \nu_D^{-1})\ln(1 + b/h)$, whose sign will be the same as the sign of $\nu_D - \nu_T$. Define the critical value

$$b_c = h\left(\frac{e^{\nu_T V} - 1}{\nu_T V}\right)^{\frac{\nu_D}{\nu_T - \nu_D}}\left(\frac{e^{\nu_D m V} - 1}{\nu_D m V}\right)^{\frac{\nu_T}{\nu_D - \nu_T}} \exp\left[\frac{\nu_D \nu_T (m-1)V}{2(\nu_T - \nu_D)}\right] - h,$$

where $\exp[x] = e^x$. Then for $b > \max\{b_T, b_D, b_c\}$, the DS policy achieves the lower inventory cost if and only if $\nu_D - \nu_T > 0$, where $\nu_{\Re}$ is the exponential parameter for the steady state distribution of the RBM associated with routing scheme $\Re$. Because $\rho_D < \rho_T$, the condition $\sigma_D^2 > \sigma_T^2$ is required for the TSP policy to be preferred. For the case of deterministic travel times it follows that $\sigma_D^2 = \sigma_T^2$, and so DS dominates in the high backorder case. Moreover, if both μ_D and μ_T are finite then the difference in mean distance travelled must be $O(n^{-1/2})$; see equation (62). Hence, in the heavy traffic limit, we actually have $\sigma_D^2 = \sigma_T^2$.

This result is somewhat counterintuitive: since the TSP policy makes smaller and more frequent deliveries to each retailer, it might be expected to outperform the DS scheme in terms of inventory cost. However, for large backorder penalties, the TSP policy sacrifices efficiency over the long run via its smaller drift, and causes the total inventory to spend too much time in the expensive backorder regions.

5 The IRP with Dynamic Routing

5.1 Formulation of the Limiting Control Problem

Consider now a situation where, once the vehicle is loaded at the warehouse, it can embark on either a full load TSP tour or a direct shipment to some retailer. All other aspects of the

problem (e.g. sources of uncertainty, cost structure) remain the same as in the fixed routing cases. Because the formulation of the dynamic problem is a natural extension of the two fixed routing problems described earlier, we only sketch the argument and refer readers to Rubio for a detailed treatment; furthermore, much of the earlier notation will be reused.

If $Q_i(0) = 0$ then the system state equations are given by

$$Q_i(t) = V_i S_T(B_T(t)) + V_i S_D(B_D(t)) - D_i(t) + \epsilon_i^D(t) + \epsilon_i^T(t) \quad \text{for} \quad t \geq 0, \tag{59}$$

and the cumulative idle time process is $I(t) = t - B_T(t) - B_D(t)$ for $t \geq 0$. Notice that, according to the previous definitions of the delivery allocation controls, $\epsilon_i^T(t) + \epsilon_i^D(t)$ represents the cumulative deviation from the nominal allocation over past TSP cycles/DS trips plus the amount delivered over the current cycle/trip at retailer i.

The first step in the development of a heavy traffic approximation is to characterize the influence of the routing control on the total netput. Let $\delta(t) = B_D(t)/(B_T(t) + B_D(t))$ denote the cumulative fraction of busy time that the DS service has been used. Then the netput for the dynamic routing IRP system is

$$\chi(t) = \left(\frac{V}{\theta_T}(1 - \delta(t)) + \frac{\lambda V}{2\sum_i \lambda_i \theta_{0i}}\delta(t) - \lambda\right) t + V\left[\mathcal{S}_T(B_T(t)) + \mathcal{S}_D(B_D(t))\right] - \mathcal{D}(t). \tag{60}$$

Notice that this equation reduces to (13) when TSP routing is always used. Summing the equations in (59) over the retailers and substituting the relevant definitions into (60), we obtain the following expression for the total inventory in terms of the netput and controls:

$$Q(t) = \chi(t) - \left(\frac{V}{\theta_T}(1 - \delta(t)) + \frac{\lambda V}{2\sum_i \lambda_i \theta_{0i}}\delta(t)\right) I(t) + \epsilon(t), \tag{61}$$

where $\epsilon(t) = \epsilon_T(t) + \epsilon_D(t)$.

The heavy traffic conditions are a union of the conditions for the two fixed routing problems. Conditions (22) and (51) require the traffic intensity under both the TSP and DS policies to approach one in the limit; however, now we do not require μ_T to be positive, because the condition $\mu_D > 0$ ensures that a stable control is possible. Hence, in terms of the problem data, the retailers are required to be located fairly close together relative to their distance to the warehouse. More precisely, the average travel times must satisfy

$$\frac{\theta_T - 2\theta_{0i}}{\theta_T} = O\left(\frac{1}{\sqrt{n}}\right) \quad \text{for all} \quad i = 1, \ldots, m. \tag{62}$$

One consequence of equation (62) is that the quantities V/θ_T and $\lambda V/(2\sum_i \lambda_i \theta_{0i})$ appearing in (61) only differ by $O(n^{-1/2})$; hence, in the heavy traffic limit, these two terms are equal and the coefficient in front of the idleness term in (61) is a constant and independent of the control policy employed. Nonetheless, we will maintain equation (61) as is, and view the time-dependent coefficient as a refinement of the heavy traffic limit.

Let $\Re(t) \in \{\text{TSP}, \text{DS}\}$ denote the routing mode that is used at time t in the limiting control problem. We want to consider intervals of time during which one of the two controls is used exclusively, and approximate the resulting normalized netput process by a diffusion with control-dependent drift and variance. In order to do this we need to reconcile the different definitions of the netput process that are approximated in the two cases: the embedded netput process in the TSP case, and the averaged netput process in the DS case. The lack of a cyclic structure in the DS case forced us to use averaging there, while the choice of embedding in the TSP case was made for convenience. Thus we translate the embedded netput process from the TSP case, $\tilde{\chi}(t)$, to an averaged netput process $\chi(t)$. This translation consists of adding a constant, $\hat{\eta}$, which needs to be determined: $\chi(t) = \tilde{\chi}(t) + \hat{\eta}$. The same constant will be used to translate the normalized total inventory process for the TSP case as well, and can be calculated in that context using the relationship between total embedded inventory x and cycle placement variables x_i developed in §2.5. Let $\bar{x}$ denote the average inventory level we are seeking. Then, since the average of the sum of the inventory levels is equal to the sum of the averages, we have $\bar{x} = \sum_i (x_i + \hat{V}_i/2) = \sum_i x_i + \hat{V}/2$. Combining this with (27) yields $\bar{x} = x - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} + \hat{V}/2$, so that

$$\hat{\eta} = \frac{\hat{V}}{2} - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} \ . \tag{63}$$

The role of this translation constant in the implementation of the proposed policy is described in the next subsection.

By the arguments used in the TSP and DS cases, we can deduce that the normalized averaged total netput process $\hat{\chi}$ is well approximated by a diffusion process $X(t, \Re(t))$ with control-dependent drift and variance given by

$$\mu(\Re(t)) = \begin{cases} \mu_D & \text{if} \quad \Re(t) = \text{DS} \\ \mu_T & \text{if} \quad \Re(t) = \text{TSP} \end{cases} \quad \text{and} \quad \sigma^2(\Re(t)) = \begin{cases} \sigma_D^2 & \text{if} \quad \Re(t) = \text{DS} \\ \sigma_T^2 & \text{if} \quad \Re(t) = \text{TSP} \end{cases} ,$$

respectively. In other words, the routing control switches the netput of the system from one Brownian motion to another. Furthermore, these two Brownian motions have the same parameters as the diffusion limits of the corresponding fixed routing cases. By equation (62), σ_D^2 and σ_T^2 only differ by $O(n^{-1/2})$; once again, we retain this refinement of the heavy traffic limit.

Because the normalized idleness process $Y(t)$ will only be exerted on a set of measure zero, we are free to choose $\Re(t)$ for all times t, not just the nonidling times. If we let $\hat{\epsilon}(t) = \epsilon(nt)/\sqrt{n}$, let $\hat{\delta}(t) = \delta(nt)$ be the fraction of busy time devoted to DS in the heavy traffic system, and define the averaged total inventory level for the system as

$$Z(t, \Re(t)) = X(t, \Re(t)) - \left(\frac{V}{\theta_T}(1 - \hat{\delta}(t)) + \frac{\lambda V}{2\sum_i \lambda_i \theta_{0i}} \hat{\delta}(t)\right) Y(t), \tag{64}$$

then we obtain the same time scale decomposition as before: under the fluid scaling, $Z(t)$ is fixed and the individual fluid levels $(\bar{W}_1(t), \ldots, \bar{W}_m(t))$ move deterministically at a finite rate. The exact evolution of $(\bar{W}_1(t), \ldots, \bar{W}_m(t))$ is determined by the averaged inventory level, the cycle placement and the routing scheme in use at time t. We may therefore decompose the dynamic routing IRP into: (i) given $Z(t) = x$ and routing mode $\Re(t) \in \{\text{TSP}, \text{DS}\}$, use the results in §2 and §3 to determine the optimal cycle placement and the corresponding inventory cost function $g_{\Re(t)}(x)$; (ii) choose the nonanticipating control $(Y(t), \Re(t))$ (where Y is nondecreasing and right continuous) to minimize

$$\limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T g_{\Re(t)}(Z(t, \Re(t)))dt - \hat{r}Y(T)\right] \tag{65}$$

subject to (64).

5.2 Optimization of a Triple Threshold Policy

The diffusion control problem (64)-(65) appears to be difficult to tackle for two reasons: the coefficient in front of the control $Y(t)$ in (64) depends upon the routing control $\Re(t)$, and the control-dependent cost function $g_\Re(x)$ is very complex. An algorithm is used in the next section to numerically compute the solution to (64)-(65). Here, we specialize our analysis to the following triple threshold policy that is characterized by the parameters $z_1 \le z_2 \le z_3$: the

vehicle is busy whenever the total averaged inventory $Z(t) < z_3$, and idles when $Z(t) \geq z_3$; while busy, the vehicle uses the TSP routing scheme whenever $Z(t) \in [z_1, z_2)$, and uses the DS mode whenever $Z(t) < z_1$ or $Z(t) \in [z_2, z_3)$. Our goal in this subsection is to find the optimal values for the parameters (z_1, z_2, z_3).

Motivated by our analysis of the stochastic ELSP with setup times in Markowitz, Reiman and Wein, we conjecture that the most general form of the optimal policy to the diffusion control problem (64)-(65) is of the triple threshold form described above. In terms of the ELSP, the DS policy corresponds to using large lot sizes and the TSP scheme corresponds to using small lot sizes: large lot sizes and the DS policy both use the server (which is the vehicle here) in a more efficient fashion. The most general form of the ELSP solution derived in Markowitz, Reiman and Wein corresponds to the triple threshold policy (see Figure 5 of that paper), and in some cases (see Figure 6 of that paper) the solution could be described with fewer thresholds. We conjecture that the IRP solution is either a triple threshold policy, a double threshold policy with $z_2 = z_3$, or a single threshold policy, which could be either the TSP policy ($z_1 = -\infty$, $z_2 = z_3$) or the DS policy ($z_1 = z_2 = z_3$ or $z_1 = z_2 = -\infty$). In §7 we describe the rationale behind our conjecture.

Our analysis of the triple threshold policy requires knowledge of the stationary distribution of the controlled reflected diffusion process and the long run expected average idleness rate. The stationary distribution $\pi(x)$ must satisfy the following system of differential equations (see Karlin and Taylor 1981):

$$\frac{\sigma^2(x)}{2}\frac{d^2}{dx^2}\pi(x) - \mu(x)\frac{d}{dx}\pi(x) = 0 \qquad x < z_3, \tag{66}$$

$$\frac{\sigma^2(x)}{2}\frac{d}{dx}\pi(x) - \mu(x)\pi(x) = 0 \qquad x = z_3, \tag{67}$$

where the diffusion parameters are given by

$$\mu(x) = \begin{cases} \mu_T & \text{if } x \in [z_1, z_2) \\ \mu_D & \text{if } x < z_1 \text{ or } x \in [z_2, z_3) \end{cases} \qquad \text{and} \qquad \sigma^2(x) = \begin{cases} \sigma_T^2 & \text{if } x \in [z_1, z_2) \\ \sigma_D^2 & \text{if } x < z_1 \text{ or } x \in [z_2, z_3) \end{cases}.$$

The only continuous density that satisfies (66) and (67) is

$$\pi(x) = \begin{cases} k_1\hat{\nu}_D e^{\hat{\nu}_D(x-z_1)} & \text{if } x \le z_1 \\ k_2\hat{\nu}_T e^{\hat{\nu}_T(x-z_2)} & \text{if } x \in (z_1, z_2] \\ k_3\hat{\nu}_D e^{\hat{\nu}_D(x-z_3)} & \text{if } x \in (z_2, z_3] \end{cases} . \qquad (68)$$

where

$$k_1 = \frac{\hat{\nu}_T}{\hat{\nu}_T(1 + e^{\hat{\nu}_T(z_2-z_1)}e^{\hat{\nu}_D(z_3-z_2)} - e^{\hat{\nu}_T(z_2-z_1)}) + \hat{\nu}_D(e^{\hat{\nu}_T(z_2-z_1)} - 1)},$$

$$k_2 = \frac{\hat{\nu}_D e^{\hat{\nu}_T(z_2-z_1)}}{\hat{\nu}_T(1 + e^{\hat{\nu}_T(z_2-z_1)}e^{\hat{\nu}_D(z_3-z_2)} - e^{\hat{\nu}_T(z_2-z_1)}) + \hat{\nu}_D(e^{\hat{\nu}_T(z_2-z_1)} - 1)} \quad \text{and}$$

$$k_3 = \frac{\hat{\nu}_T e^{\hat{\nu}_T(z_2-z_1)}e^{\hat{\nu}_D(z_3-z_2)}}{\hat{\nu}_T(1 + e^{\hat{\nu}_T(z_2-z_1)}e^{\hat{\nu}_D(z_3-z_2)} - e^{\hat{\nu}_T(z_2-z_1)}) + \hat{\nu}_D(e^{\hat{\nu}_T(z_2-z_1)} - 1)};$$

this is the stationary density for the total inventory process.

Now we turn to the expected idleness rate. Taking expectations on both sides of (64), rearranging terms, dividing through by t and taking the limit as $t \to \infty$ we have that

$$\lim_{t\to\infty} \frac{1}{t} E\left[\left((1-\hat{\delta}(t))\frac{V}{\theta_T} + \hat{\delta}(t)\frac{\lambda V}{2\sum_i \lambda_i\theta_{0i}}\right) Y(t)\right] =$$

$$\lim_{t\to\infty} \frac{1}{t} E[X(t, \Re(t))] - \lim_{t\to\infty} \frac{1}{t} E[Z(t, \Re(t))]. \qquad (69)$$

The first term on the RHS of (69) corresponds to the asymptotic growth rate of the netput process under the proposed policy. If we let $\delta = lim_{t\to\infty}\hat{\delta}(t)$ denote the long run fraction of time that the vehicle uses the DS routing policy, then the long run growth rate is

$$\lim_{t\to\infty} \frac{1}{t} E[X(t, \Re(t))] = (1-\delta)\mu_T + \delta\mu_D. \qquad (70)$$

By the definition of the triple threshold policy, it follows that

$$\delta = \int_{-\infty}^{z_1} \pi(x)dx + \int_{z_2}^{z_3} \pi(x)dx = k_3 + k_1(1 - e^{\hat{\nu}_D(z_2-z_3)}).$$

The left side of (69) can be expressed as

$$\left((1-\delta)\frac{V}{\theta_T} + \delta\frac{\lambda V}{2\sum_i \lambda_i\theta_{0i}}\right) \lim_{t\to\infty} \frac{1}{t} E\left[Y(t)\right]. \qquad (71)$$

Finally, according to the steady state distribution for $Z(t)$ in (68), $\bar{z} = \lim_{t\to\infty} E[Z(t)]$ exists and is finite. Hence, the second term on the right side of (69) vanishes in the limit. Canceling this term and substituting (70) and (71) into (69) gives

$$\bar{y} = \lim_{t\to\infty} \frac{1}{t} E[Y(t)] = \sqrt{n}\left(1 - \frac{\rho_T(1-\delta)2\sum_i \lambda_i\theta_{0i} + \rho_D\delta\lambda\theta_T}{(1-\delta)2\sum_i \lambda_i\theta_{0i} + \delta\lambda\theta_T}\right). \tag{72}$$

We may now write an expression for the steady state cost of the IRP under the triple threshold dynamic routing policy as a function of the control parameters. The problem is hence reduced to finding (z_1^*, z_2^*, z_3^*) that minimize

$$\hat{F}(z_1, z_2, z_3) = \int_{-\infty}^{z_1} g_D(x)\pi(x)dx + \int_{z_1}^{z_2} g_T(x)\pi(x)dx + \int_{z_2}^{z_3} g_D(x)\pi(x)dx - \hat{r}\bar{y}. \tag{73}$$

Unfortunately, $\hat{F}(z_1, z_2, z_3)$ is rather complicated and a closed form solution for the optimal control parameters does not seem possible. Furthermore, an explicit expression for $\hat{F}(z_1, z_2, z_3)$ is not available in general because its exact form depends on the relationship between $(\hat{\alpha}_T, \hat{\beta}_T)$ and $(\hat{\alpha}_D, \hat{\beta}_D)$, which are the parameters that define the three characteristic regions for the cycle placement and inventory cost solutions. By considering different problem parameters, one may get either $\hat{\beta}_D > \hat{\beta}_T$ or $\hat{\beta}_D < \hat{\beta}_T$.

5.3 The Proposed Policy

The first step in mapping the solution of the diffusion control problem into a policy for the dynamic IRP is to obtain unscaled threshold levels. If we define the unscaled thresholds as $w_i = \sqrt{n}\, z_i$ and make the parameter scaling substitutions as in the fixed routing cases, then the scaling factor n cancels out of the expression for

$$F(w_1, w_2, w_3) = \sqrt{n}\,\hat{F}\left(\frac{w_1}{\sqrt{n}}, \frac{w_2}{\sqrt{n}}, \frac{w_3}{\sqrt{n}}\right), \tag{74}$$

and what remains is a function only of the original system parameters. In §6.2, the thresholds minimizing (74) are determined numerically, and we refer to these cost-minimizing thresholds as (w_1^*, w_2^*, w_3^*).

We now describe how these thresholds dictate whether the vehicle employs the TSP mode, the DS mode or idles; this decision is made at the epochs when the vehicle is at

the warehouse. Because of the translation of the embedded total inventory process into the averaged total inventory process in (63), we must keep track of the most recently employed shipping option. If DS was most recently used then we calculate the unscaled averaged inventory q in (55) and compare it to the unscaled threshold levels (w_1^*, w_2^*, w_3^*) to determine the TSP/DS/idle policy. If TSP was most recently used then we define q by combining (35) and (63):

$$q = \sum_j Q_j(t_i^-) - \eta_i + \sum_j \lambda_j \theta_{0j}^{TSP} + \eta, \tag{75}$$

where

$$\eta = \sqrt{n}\hat{\eta} = \frac{V}{2} - \sum_i \lambda_i \theta_{0i}^{TSP} . \tag{76}$$

is the unscaled translation constant. In this case the TSP/DS/idling decision is determined by comparing q in (75) to the unscaled threshold levels (w_1^*, w_2^*, w_3^*). Finally, the detailed allocation decisions (which retailer to visit next in the DS case and how many units to deliver to each retailer in the TSP case) are determined exactly as in Sections 2 and 3, except that we now use the value of q in (75) in the TSP case.

6 Computational Results

This section contains a series of computational experiments aimed at assessing the accuracy of the heavy traffic analysis and determining what aspects of the control policy are most important for good system performance. The computational study is described in two parts: the fixed routing IRP and the dynamic routing IRP.

6.1 Fixed Routing IRP

The Monte Carlo simulation experiments performed in this subsection consider systems that have five retailers and Poisson demand processes. We also set the transportation cost rate r equal to zero and concentrate on the inventory cost. The total arrival rate λ is varied to obtain different utilization rates; however, the fraction of demand represented by retailer i is fixed so that $\lambda_1 = \lambda/5, \lambda_2 = \lambda/10, \lambda_3 = \lambda/10, \lambda_4 = \lambda/5$ and $\lambda_5 = 2\lambda/5$. The travel time

random variables T_{ij} are iid second order Erlang. The mean travel times are adjusted so that $10\theta_T = V$ always holds; this allows us to consider several vehicle sizes while maintaining the traffic intensity at $\rho_T = 0.1\lambda$.

We perform four simulation experiments aimed at various aspects of system performance.

Experiment 1: The first set of simulation runs quantifies the cost improvement obtained under the TSP policy by recalculating the cycle placement at each retailer, as opposed to determining these values only once per cycle (i.e., when the vehicle is at the warehouse). We let $b_i = b = 5$, $h_i = h = 1$ for $i = 1, \ldots, 5$ and consider the mean travel times $\theta_{01} = \theta_{50} = \theta_T/10$, $\theta_{12} = \theta_{23} = \theta_{34} = \theta_{45} = \theta_T/5$. These travel times are consistent with a *pentagon* structure, where the five retailers are placed at the vertices of a pentagon, and the warehouse is located midway between stations 1 and 5.

We consider nine different scenarios, which are generated by the different combinations of three vehicle sizes $(100, 10, 5)$ and three traffic intensities $(0.5, 0.7, 0.9)$; notice that some of these scenarios grossly violate the heavy traffic conditions. For all cases with $\rho_T \leq 0.7$, we simulated three replications of 36,000 time units (starting with an empty system and discarding the first 2000 time units) with cycle placement recalculation at the retailers and three more with calculation only at the warehouse; for the $\rho_T = 0.9$ instances, the length of each replication was increased to 240,000 time units (discarding the first 20,000 time units). This simulation design was used throughout our study and allowed us to keep the standard deviation of the cost estimate under 1% of its mean.

Table 1 summarizes the results of the experiment. The entries in the table represent the increase in the average inventory cost when delivery sizes are calculated only at the warehouse, and not adjusted over the course of the tour. As predicted by heavy traffic theory, the advantage obtained by recalculation at the retailers becomes quite small when the traffic intensity is high (about 1% when $\rho_T = 0.9$). However, the recalculation advantage increases with small vehicle sizes at lower traffic intensities. These observations complement those in Kumar, Schwarz and Ward, who focus on this issue (calculating delivery allocations once per cycle or at each retailer) using a much different model. All subsequent TSP simulations

	$\rho_T = 0.5$	$\rho_T = 0.7$	$\rho_T = 0.9$
$V = 100$	4.0%	5.0%	1.1%
$V = 10$	7.1%	18.6%	0.7%
$V = 5$	7.3%	18.1%	1.1%

Table 1: Cost increase when placement calculation is only at warehouse.

employ the cycle placement recalculation at the retailers.

Experiment 2. The second simulation experiment assesses the accuracy of the heavy traffic analysis by comparing the cost incurred under the derived base stock levels with the cost incurred under the best possible base stock level. We maintain the same set-up as in the first experiment, except that asymmetric cost cases are also considered, where the holding rates are $(1, 1, 2, 2, 2)$ and the backorder rates are $(5, 10, 5, 10, 5)$ for the five retailers, respectively. For each of these 18 cases (three vehicle sizes, three traffic intensities and two cost structures), we performed an exhaustive search in a series of simulations (each consisting of three replications with the length described in Experiment 1) to find the base stock level that provides the lowest system cost.

Table 2 summarizes the results; each entry represents the suboptimality (within the class of base stock policies) of the cost incurred by using the derived base stock level instead of the optimal base stock level found by exhaustive search. The base stock levels derived from the heavy traffic analysis are very accurate for moderate and high traffic intensities, and only degrade in the $\rho_T = 0.5$, small vehicle size scenarios, which are not apt to occur in practice.

Experiment 3. Now we study the performance of the direct shipping policy, and compare it to the performance of the TSP policy on the same system. As before, this is done by comparing the average cost obtained under the derived base stock levels with that under the optimal base stock level found by exhaustive search. Because the DS policy has a huge drift advantage over the TSP policy in the pentagon topology used for experiments 1 and 2, this experiment uses the travel times $\theta_{0i} = 0.45\theta_T$ for $i = 1, \ldots, 5$ and $\theta_{12} = \theta_{23} =$

		$\rho_T = 0.5$	$\rho_T = 0.7$	$\rho_T = 0.9$
$V = 100$	Symm.	0.0%	0.9%	2.6%
	Asym.	0.6%	2.2%	0.0%
$V = 10$	Symm.	19.1%	4.3%	1.7%
	Asym.	6.7%	2.2%	1.8%
$V = 5$	Symm.	14.6%	1.1%	1.5%
	Asym.	11.6%	1.1%	0.6%

Table 2: Cost suboptimality of derived base stocks for pentagon TSP.

$\theta_{34} = \theta_{45} = \theta_T/40$, so that $\rho_T = 0.1\lambda$ and $\rho_D = 0.09\lambda$. This case will be referred to as the *wedge* topology, since these travel times are consistent with such a shape. In practice, TSP tours are often generated by placing the warehouse at the center of the "pie", dividing the pie into wedges and solving a TSP on each wedge; see Figure 1 of Bell et al. and Figure 3 of Federgruen and Simchi-Levi. The other problem parameters remain as in the symmetric cost scenarios in Experiment 2, except for the fact that we also simulate the DS policy for the case when $\lambda = 10$. The TSP policy is not simulated for this case because it corresponds to $\rho_T = 1$, and the system is not stable under this scheme. Hence, we consider 12 cases (four traffic intensities and three vehicle sizes) for the DS policy and nine cases for the TSP.

Tables 3 and 4 summarize the results of this experiment. The entries in Table 3 compare the performance of the proposed base stock level to the cost obtained under the best base stock level for *the same policy*. The results for the TSP policy are roughly comparable to the corresponding results for the pentagon topology in Table 2, although the base stock levels in Table 3 do not seem to be as accurate for the $\rho_T = 0.9$ cases. The derived base stock levels for the DS case are very accurate, even when the heavy traffic conditions are severely violated.

Table 4 presents a comparison of the average inventory cost for the DS and TSP policies. The percentage difference between the TSP cost and the DS cost is given by

$$\frac{\text{TSP cost} - \text{DS cost}}{\text{DS cost}} \times 100\%.$$

The entries labeled 'Sim.' represent the percentage difference in inventory cost when base

		$\rho_T = 0.50$ $\rho_D = 0.45$	$\rho_T = 0.70$ $\rho_D = 0.63$	$\rho_T = 0.90$ $\rho_D = 0.81$	$\rho_T = 1.00$ $\rho_D = 0.90$
$V = 100$	TSP	2.41%	5.31%	6.10%	N.A.
	DS	1.13%	2.12%	1.74%	0.52%
$V = 10$	TSP	11.04%	0.00%	3.67%	N.A.
	DS	4.30%	3.01%	2.13%	0.08%
$V = 5$	TSP	17.48%	0.00%	1.19%	N.A.
	DS	3.70%	1.75%	1.40%	0.00%

Table 3: Cost suboptimality of derived base stocks for wedge topology.

stock levels are found by exhaustive search. Notice that for low utilization levels and large vehicle sizes the TSP policy enjoys a considerable advantage over the DS scheme. This advantage erodes as the traffic intensity increases until, for the cases where $\rho_T = 1$, the TSP cost becomes unbounded and the DS policy is trivially preferred. Recall that the percentage differences in Table 4 only assess the inventory costs, and the DS policy will always incur lower transportation costs than the TSP policy. Hence, the desired policy is a function of the transportation cost rate r.

The entries labeled 'Pred.' represent the difference in inventory costs as predicted by the heavy traffic analysis. Table 4 shows that the heavy traffic analysis provides reliable estimates for the *relative* performance of the two policies, except when the vehicle size is very small. In the one case where the prediction errs in the sign of the percentage difference ($\lambda = 9$, $V = 10$), the costs for both policies are very close. The predicted differences are larger than the simulated differences in 6 of the 9 cases, including the ($\rho_T = 0.9$, $V = 100$) heavy traffic case; this discrepancy may be due to the fact that our DS estimates are based on the idealized fluid cycles, and hence should underestimate the true heavy traffic cost under DS.

Experiment 4. The last experiment in this subsection measures the increase in cost incurred by using a base stock level different from the one proposed in the heavy traffic analysis. We already have the required data for this analysis from the exhaustive search

		$\rho_T = 0.50$ $\rho_D = 0.45$	$\rho_T = 0.70$ $\rho_D = 0.63$	$\rho_T = 0.90$ $\rho_D = 0.81$	$\rho_T = 1.00$ $\rho_D = 0.90$
$V = 100$	Sim.	-80.2%	-72.5%	-29.1%	N.A.
	Pred.	-78.4%	-71.8%	-22.1%	∞
$V = 10$	Sim.	-66.5%	-58.6%	-3.0%	N.A.
	Pred.	-76.7%	-65.0%	2.5%	∞
$V = 5$	Sim.	-56.7%	-64.5%	12.6%	N.A.
	Pred.	-74.3%	-57.8%	25.2%	∞

Table 4: Inventory cost comparison $\frac{\text{TSP-DS}}{\text{DS}}$: wedge topology.

performed in the simulation experiments above. Figure 2 plots three examples of the cost increase with respect to the proposed policy, as a function of the base stock level (expressed in units of vehicle size). These three cases correspond to the DS system on the wedge topology for $V = 100$ and $\lambda \in \{5, 7, 9\}$. The behavior illustrated here is typical of all other instances analyzed in our simulation experiments. Three characteristics worth noting are: (i) the inventory cost is convex in the base stock level; (ii) the cost performance remains relatively constant over a range of approximately one vehicle size around the optimal base stock level; and (iii) once the base stock level moves beyond this range in either direction the cost performance deteriorates rapidly.

6.2 Dynamic IRP

An Algorithmic Solution. In an attempt to understand the nature of the optimal solution to the dynamic IRP, we pursue a computational approach to problem (64)-(65). The algorithmic procedure, which was pioneered by Kushner (1977), approximates the diffusion process by a discrete time and space Markov chain, and then numerically solves the control problem by dynamic programming. Weak convergence methods have been developed to verify that the controlled Markov chain and its optimal cost approximate arbitrarily closely (at an increased computational expense) the controlled diffusion process and its optimal cost.

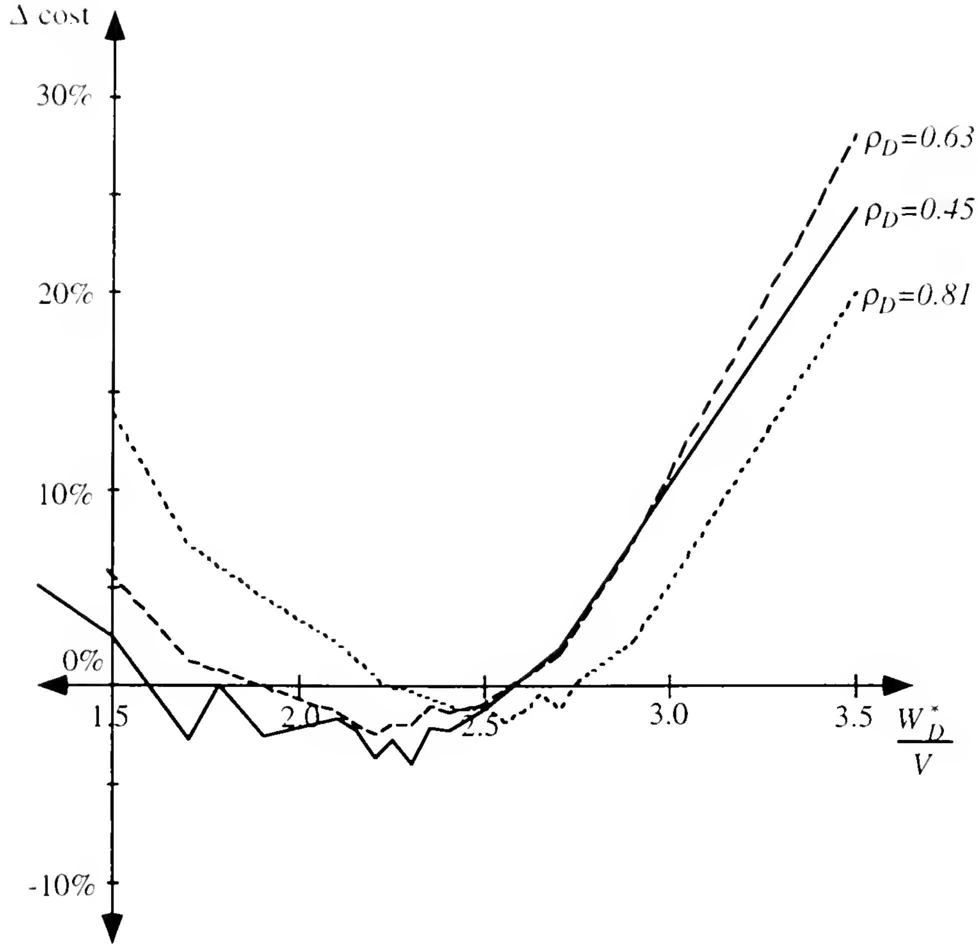

Figure 2: Sensitivity of inventory cost to base stock level.

$\lambda = 9.0$			Total Averaged Inventory
$V = 50$	$r = 500$	Idle at	101
		Use DS	$(61, 101]$
		Use TSP	$(-20, 61]$
		Use DS	$(-\infty, -20]$
		δ^*	65.0%
$V = 10$	$r = 100$	Idle at	23
		Use DS	$(11, 23]$
		Use TSP	$(-4, 11]$
		Use DS	$(-\infty, -4]$
		δ^*	73.7%
	$r = 50$	Idle at	21
		Use DS	$(15, 21]$
		Use TSP	$(-5, 15]$
		Use DS	$(-\infty, -5]$
		δ^*	46.2%

Table 5: Triple threshold policies obtained from the Markov chain approximation.

Interested readers are referred to Kushner and Dupuis (1992) for a recent account of this research area.

We developed an implementation of this algorithm and solved (64)-(65) for 36 cases that use the wedge topology and the symmetric cost structure described in experiment 3; these 36 cases, which will be enumerated later, are characterized by the vehicle capacity V, the demand rate λ and the transportation cost r. For brevity's sake, we omit a description of the algorithm and refer readers to Rubio for a detailed account of this work. (There is an error in Rubio's *description* of the implementation of the algorithm; in the improvement stage of the policy improvement algorithm on page 119, the new routing scheme $\Re_{k+1}(x)$ is found by minimizing the gain γ^ϵ in (5.71) while keeping z and $\Re(y)$ fixed for all other states $y \neq x$, *not* by minimizing the right side of (5.68) for the given $(V_k^\epsilon(x), \gamma_k^\epsilon)$.)

The numerical results are consistent with our conjectures about the optimal solution

to (64)-(65). In three of the 36 cases the policy generated by the Markov chain approximation algorithm is of the triple threshold form; in the remaining 33 cases, the solution is a degenerate form of the triple threshold policy that can be described by one or two thresholds. Table 5 specifies the proposed triple threshold policies (in terms of the unscaled inventory) for the three cases, along with the proportion of time that DS is used. As expected, the TSP scheme is used in an interval containing zero.

The Analytical Double Threshold Policy. Because the numerical results in 33 of the 36 cases can be described by one or two thresholds, and because the Markov chain procedure only offers an approximate solution to problem (64)-(65) (in particular, the solution is not independent of the heavy traffic parameter n), we turn to an analytical derivation of the double threshold policy, where the vehicle uses DS when $Z(t) < z_1$, uses TSP routing when $Z(t) \in [z_1, z_3)$ and idles when $Z(t) \geq z_3$. The parameters z_1^* and z_3^* are derived by setting z_2 equal to z_3 in the results in §5.2; in particular, the function F in (74) must be minimized. For the symmetric cost, wedge topology cases, this function can take eight different possible functional forms, and a steepest descent method was used to find the global minimum of this function; see Appendix B of Rubio for complete details.

The results for the 36 cases are presented in Table 6. The derived double threshold policy is degenerate in 22 of the 36 cases, with the TSP policy optimal in 16 cases and DS optimal in 6 cases. In the remaining 14 cases where two thresholds are required, the value of δ^* is often very close to zero, suggesting that the TSP policy is close to optimal. While results are not reported here, similar findings continued to hold for different values of the backordering-to-holding cost ratio b/h, as well as for wider or narrower wedges (i.e., for other values of θ_{01}/θ_{12}).

For the 33 cases where the Markov chain procedure generated a single or double threshold solution, these solutions matched the analytical solutions in Table 6 very closely; hence, the Markov chain solutions are not included here. For the three cases where the Markov chain procedure generated a triple threshold policy, we used simulation to compare the Markov chain policy in Table 5 to the analytically derived policy in Table 6. Table 7 shows that the double threshold solution outperforms the policy generated by the Markov chain

			$\lambda = 5.0$	$\lambda = 7.0$	$\lambda = 8.0$	$\lambda = 9.0$
$V = 100$	$r = 500$	w_3^*	20	45	75	138
		w_1^*	-312	-151	-100	-51
		δ^*	0.0%	0.0%	1.3%	8.5%
	$r = 100$	w_3^*	20	45	76	141
		w_1^*	-385	-168	-111	-58
		δ^*	0.0%	0.0%	1.0%	7.7%
	$r = 50$	w_3^*	20	45	76	142
		w_1^*	-404	-171	-112	-59
		δ^*	0.0%	0.0%	1.0%	7.5%
$V = 50$	$r = 500$	w_3^*	10	23	39	112
		w_1^*	-10^6	-64	-43	112
		δ^*	0.0%	0.0%	2.0%	100.0%
	$r = 100$	w_3^*	10	24	39	73
		w_1^*	-202	-80	-53	-28
		δ^*	0.0%	0.0%	1.3%	8.2%
	$r = 50$	w_3^*	10	24	39	73
		w_1^*	-202	-82	-54	-29
		δ^*	0.0%	0.0%	1.2%	8.0%
$V = 10$	$r = 500$	w_3^*	18	20	21	24
		w_1^*	18	20	21	24
		δ^*	100.0%	100.0%	100.0%	100.0%
	$r = 100$	w_3^*	2	6	10	24
		w_1^*	-24	-11	-8	24
		δ^*	0.0%	0.0%	3.9%	100.0%
	$r = 50$	w_3^*	2	6	10	18
		w_1^*	-31	-13	-9	-4
		δ^*	0.0%	0.0%	3.3%	11.9%

Table 6: The analytical double threshold policy.

$r =$	500	100	50
$V =$	50	10	10
$\lambda =$	9.0	9.0	9.0
(A): Double Threshold	560.9	116.9	72.5
(B): Markov Chain Appr.	596.6	124.3	85.6
$\frac{B - A}{A} 100\%$	6.4%	6.3%	18.1%

Table 7: Cost performance of Markov approximation method vs. double threshold policy.

approximation in all three cases. Although not reported here, the predicted improvement in gain from the Markov chain solution versus the derived double threshold policy was indeed relatively small for all three cases. Moreover, heavy traffic conditions (20) and (24) are violated in the 18.1% suboptimality case.

Simulation Study. With the derived double threshold policy in hand, we now perform a series of simulation runs for the dynamic routing IRP to gauge the accuracy of our heavy traffic approximations over a range of values for the problem parameters. We set $r = 500$ and consider a total of six cases (3 traffic intensities and 2 vehicle sizes), using the same five-retailer wedge system that was presented earlier. In our simulation runs, we inadvertently set η equal to zero in equations (75)-(76). For our examples, $\eta = V(40 - 41\rho_T)/80$, which equals $0.04V$, $0.09V$ and $0.14V$ when ρ_T equals 0.7, 0.8 and 0.9, respectively. Given our analysis in Figure 2, where system cost remains relatively constant over a range of approximately V around the optimal base stock level, this omission is inconsequential.

The entries in Table 8 represent the cost increase incurred by using the analytically derived double threshold policy, the best (i.e., base stock level found by exhaustive search) TSP policy or the best DS policy instead of the best double threshold policy found by exhaustive search over the (w_1, w_3) plane. In all cases, the delivery allocation is determined by the dynamic rule derived from the heavy traffic optimal cycle placement. The derived double threshold policy performs very well, and does not seem to deteriorate at lower traffic intensities or smaller vehicle sizes.

		$\lambda = 7.0$	$\lambda = 8.0$	$\lambda = 9.0$
$V = 100$	Prop.	2.3%	2.2%	2.2%
	TSP*	1.4%	2.0%	6.1%
	DS*	42.5%	27.6%	10.6%
$V = 50$	Prop.	0.8%	1.3%	3.8%
	TSP*	0.3%	0.8%	0.9%
	DS*	18.4%	11.8%	3.3%

Table 8: Suboptimality of derived double threshold and fixed routing policies.

A glance at the appropriate entries in Table 6 shows that in 5 out of the 6 cases in Table 8 the value of δ is close to either 0 or 1. The exception is the $(\lambda = 9, V = 100)$ case, where $\delta = 8.5\%$. Table 8 confirms that the best double threshold policy outperforms either of the static routing schemes in this case. We changed λ while leaving everything else fixed in an attempt to find a case where the advantage of the dynamic policy would be more dramatic. As it turns out, we could not significantly improve over the $(\lambda = 9, V = 100)$ case; in the end, the analytically derived double threshold policy was at most 5% better than the best fixed routing policy. Furthermore, in the cases where the proposed double threshold policy coincides with either of the fixed routing schemes, the cost increase incurred by choosing the wrong fixed routing scheme is quite significant (higher than 10% in all 5 cases, and up to 43% for $\lambda = 7, V = 100$). These numbers suggest that, while finding the best fixed routing scheme is very important, the advantage obtained from dynamic routing is quite small in most problem instances.

7 Summary and Conclusions

The IRP is one of the more challenging problems in operations research, especially when considered from a dynamic and stochastic viewpoint. We focus on the operational aspects of the problem and consider a system with a single capacitated vehicle that operates out of a single warehouse and services a finite set of retailers. By restricting an outgoing vehicle

to deliver full loads to either a single retailer (direct shipping or DS) or along a prespecified (TSP) tour, we avoid the combinatorial complexities inherent in the problem and maintain a sharp focus on the crucial tradeoff between inventory costs and transportation costs that lies at the heart of the IRP. Our modeling of the dynamic stochastic IRP as a queueing control problem offers a new perspective on the problem: rather than view the IRP as a variant of the vehicle routing problem, we see it as a variant of a production/inventory control problem (where the capacitated vehicle plays the role of the production system); as such, this paper is a natural descendant of Wein and Markowitz, Reiman and Wein, which consider more conventional production/inventory control problems.

By assuming that the system is operating in the (suitably defined) heavy traffic regime, we approximate the queueing control problem by a diffusion control problem. When only TSP tours are allowed, this modeling approach, together with the application of a heavy traffic time scale decomposition, allows us to fully characterize the solution to the diffusion control problem, thereby generating an operating policy for the original system. By assuming the existence of a fixed sequence of retailer visits than can achieve constant inter-delivery times to each retailer in the fluid limit, we perform a similar analysis for the DS case. The control policy in both cases is characterized by a vehicle idling policy, which dictates whether a vehicle at the warehouse should sit idle or set out with a full load, and a dynamic allocation policy, which specifies how many units to leave off at each retailer under a TSP scheme, and which retailer to visit next in the DS scheme.

We also consider the case where dynamic route selection (either TSP or DS) is allowed. The diffusion control problem is solved numerically and a class of triple threshold policies is analyzed. Finally, a series of simulation studies is performed to complement the heavy traffic analysis.

Our key findings can be summarized as follows:

- The inventory component of the total long run average cost depends on the stochastic characteristics of the system, while the transportation component for a fixed routing scheme is determined solely from first moment information.

- The vehicle idling policy is characterized by an aggregate base stock level: the vehicle idles at the warehouse whenever the total retailer inventory exceeds a certain threshold level. The value of the optimal base stock level is independent of the transportation cost rate. Although the existing IRP literature does not typically address the vehicle idling issue, our simulation results show that system performance is quite sensitive to the value of the base stock level, deteriorating rapidly when the base stock level differs from the optimal value by more than the vehicle capacity. Moreover, simulation results also confirm that the system cost under our derived base stock levels are typically within several percent of the cost achieved by the best (found by exhaustive search using simulation) base stock level, unless the heavy traffic conditions are grossly violated (e.g., traffic intensity equals 0.5 and vehicle capacity is less than or equal to 10 units).

- The allocation of load among the retailers is dictated by the desire to concentrate most of the total inventory (backorders) at the site with the smallest holding (backorder) cost rate.

- Dynamic (i.e., state-dependent or closed-loop) delivery allocations greatly outperform their static (state-independent or open-loop) counterparts in a stochastic environment. In fact, central limit theorem arguments indicate that static delivery allocations cause the absolute value of the inventory levels to grow as the square root of time, thereby leading to unbounded costs over the long run.

- The relative advantage of recalculating the load allocation at each retailer within a TSP tour, as opposed to setting it once at the beginning of each tour, decreases as utilization increases, and vanishes in the heavy traffic limit.

- The policy that achieves the lowest transportation cost is the one that delivers the largest amount per unit time travelled (subject to meeting average demand). Therefore, direct shipping is the most transportation-efficient routing scheme; although this fact is a trivial consequence of the triangle inequality, it is perhaps our most important observation. This fact helps highlight the basic cost tradeoff in the IRP: *DS leads to*

smaller transportation costs, but TSP routing, by making smaller and more frequent deliveries, may lead to smaller inventory costs. This result also implies that *DS has a larger stability region than TSP*; that is, for any given problem instance, one can increase the demand rates to a level where the TSP routing scheme has a traffic intensity greater than or equal to one and the DS scheme has an intensity less than one. Moreover, our analysis highlights the danger in employing a myopic policy that always minimizes current inventory costs (for example, have each outgoing vehicle satisfy as many backorders as possible, regardless of location); such a policy, which is similar in spirit (and consequence) to a production lot-sizing policy that frequently breaks a setup in order to satisfy backordered demand, can easily become unstable.

- Heavy traffic analysis shows that for cost-symmetric systems with sufficiently high backorder costs, DS will be preferred to TSP routing.

- Simulation results show that there is often a large difference in performance between the DS and TSP policies. Although the traffic intensity, backorder costs and transportation costs all play a significant role, the topology probably plays the largest role in the relative attractiveness of each policy. For systems with relatively high loads, it appears that TSP could only be a desirable alternative when the tour is wedge-shaped, as is often the case in practice.

- The heavy traffic analysis accurately predicts the relative cost of using the fixed DS or fixed TSP schemes. Hence, our procedure can be used as an aid in higher level decisions, as discussed at the end of this paper.

- If one can dynamically choose between the DS and TSP options, we conjecture that the most general form of the solution is a triple threshold policy characterized by $w_1 \leq w_2 \leq w_3$: if the total retailer inventory is less than w_1 then DS is preferred, if it is in the interval $[w_1, w_2)$ then TSP is preferred and if it is in the interval $[w_2, w_3)$, where w_3 is the idling threshold, then DS is preferred. Our rationale is as follows: if the absolute value of the total retailer inventory is large then the routing scheme may

have little effect on the rate at which inventory costs are incurred. In these cases, DS may be preferable because it incurs smaller transportation costs. In addition, the efficiency of DS has a tendency to increase the total inventory level relative to TSP (in the diffusion control problem the DS option has a larger drift than the TSP option), and so DS will be even more attractive when the total inventory is much less than zero, as it will help to decrease future backorders. However, when the total inventory is in the interval $[w_1, w_2)$, which should contain zero in the nondegenerate case, the frequent deliveries of TSP lead to less backorders and smaller inventory costs, making it the more attractive alternative. Finally, because the effective penalty for using the inefficient TSP policy decreases when the total inventory is large, we believe that in most cases the optimal solution will be no more complex than a double threshold policy, where $w_2 = w_3$. This state of affairs is somewhat analogous to the stochastic ELSP with setup times analyzed in Markowitz, Reiman and Wein, where large (small) lot sizes correspond to DS (TSP).

- We computed the numerical solution to the diffusion control problem corresponding to the dynamic IRP for a number of instances, and the results were consistent with our conjectures: the most general optimal policy was of the triple threshold form, and in most cases a degenerate form of the policy was optimal: either the fixed DS case ($w_1 = w_2 = -\infty$ or $w_1 = w_2 = w_3$), the fixed TSP case ($w_1 = -\infty$, $w_2 = w_3$) or the double threshold policy ($w_2 = w_3$). Moreover, in our limited simulation experiments, we did not find a numerically computed triple threshold policy that outperformed the analytically derived double threshold policy (although we did not search beyond the computed triple threshold values). By performing many exhaustive searches using simulation, we also found that the best double threshold policy differed from the better of the two fixed routing policies in only a narrow range of system parameter space; in this range, the best TSP and DS policies achieve fairly similar performance. Hence, coupling this observation with a previous one suggests that *finding the best fixed route policy is very important while allowing for dynamic routing provides a much*

less substantial benefit; this is particularly true in light of the increased complexity of implementing a dynamic routing scheme.

In summary, the important operational levers for the IRP include the aggregate base stock level, the dynamic allocation policy and the choice of fixed routing scheme, but not the dynamic routing policy. Moreover, these key decisions are interrelated and a unified stochastic control model, such as the one considered here, is required for achieving reliable system performance.

Two topics for future research naturally come to mind. The first is to extend the dynamic routing scheme so as to allow K different types of routes (where $K > 2$) and/or to consider cyclic routes that use a combination of DS and TSP (e.g., a cycle could consist of a TSP tour through retailers 1, 2 and 3, followed by a direct shipment to retailer 2). Although in theory these extensions could be incorporated and system improvements could be achieved, the analysis would be tedious and it is doubtful that any additional insights would be found.

Perhaps the most fruitful area for future research would be to develop the necessary steps for a hierarchical approach to the general (multi-vehicle, multi-warehouse) IRP; such an approach would be similar in spirit to the vehicle routing analysis performed by Simchi-Levi, but would also incorporate the inventory cost component. Our results for fixed route policies provide estimates for the operating cost for any system given a particular assignment of retailers to vehicles and vehicles to warehouses. Motivated by our observation that the best fixed route policy performs nearly as well as the best dynamic policy over a broad range of parameters, the first level up in the hierarchy could implement some interexchange optimization algorithm (e.g. a k-opt algorithm as used in the deterministic vehicle routing literature) to find the best such route. Higher levels in the hierarchy could then be used to select the best possible assignment of vehicles and retailers, and the total number of vehicles to have in the system. At an even higher level, these results could be used to decide on the number and location of warehouses.

Acknowledgment

This research was supported by a grant from the Leaders for Manufacturing Program at MIT, NSF grant DDM-9057297 and EPSRC grant GR/J71786. The last author thanks the Statistical Laboratory at the University of Cambridge for its hospitality while some of this work was carried out.

References

Ackoff, R. L. 1987. OR, A Post Mortem. *Operations Research* **35**, 471-474.

Bell, W., L. M. Dalberto, M. Fisher, A. Greenfield, R. Jaikumar, P. Kedia, R. G. Macj and P. J. Prutzman. Improving the Distribution of Industrial Gases with an On-Line Computerized Routing and Scheduling Optimizer. *Interfaces* **13**, 4-23.

Billingsley, P. 1968. *Convergence of Probability Measures*. John Wiley and Sons, New York.

Chan, L. M. A., A. Federgruen and D. Simchi-Levi. 1994. Probabilistic Analyses and Practical Algorithms for Inventory-Routing Models. Graduate School of Business, Columbia University.

Coffman, E. G. Jr., A. A. Puhalskii and M. I. Reiman. 1995a. Polling Systems with Zero Switchover Times: A Heavy-Traffic Averaging Principle. *Annals of Applied Probability* **5**, 681-719.

Coffman, E.G., A. A. Puhalskii and M. I. Reiman. 1995b. Polling Systems with Switchover Times: A Heavy-Traffic Averaging Principle. AT&T Bell Laboratories, Murray Hill, NJ.

Dror, M. and M. Ball. 1987. Inventory Routing: Reduction from an Annual to Short Period Problem. *Naval. Res. Log. Quarterly* **34**, 891-905.

Federgruen, A. and Z. Katalan. 1994. The Stochastic Economic Lot Scheduling Problem: Cyclic Base-Stock Policies with Idle Times. Graduate School of Business, Columbia University, New York, NY. To appear in *Management Science*.

Federgruen, A. and D. Simchi-Levi. 1992. Analytical Analysis of Vehicle Routing and Inventory-Routing Problems. *Handbooks in Operations Research an Management Science*,

the volume on Networks and Distribution. M. Ball, T. Magnanti, C. Monma and G. Nemhauser, eds.

Federgruen, A. and P. Zipkin. 1984. A Combined Vehicle Routing and Inventory Allocation Problem. *Operations Research.* **32**, 1019-1037.

Golden, B.L. and A. A. Assad (Editors). 1988. *Vehicle Routing: Methods and Studies.* North-Holland Publishers.

Harrison, J.M. 1988. Brownian Models of Queueing Networks with Heterogeneous Customer Populations. In *Stochastic Differential Systems, Stochastic Control Theory and Applications*, IMA Volume **10**, W. Fleming and P.L. Lions (eds.). Springer-Verlag, New York, 147-186.

Harrison, J.M. 1985. *Brownian Motion and Stochastic Flow Systems.* John Wiley and Sons, New York.

Iglehart, D. L. and W. Whitt. 1970. Multiple Channel Queues in Heavy Traffic I. *Advances in Applied Probability.* **2**, 150-177.

Karlin S. and H. M. Taylor. 1981. *A Second Course in Stochastic Processes.* Academic Press, Inc. San Diego, CA.

Kumar, A., L. B. Schwarz and J. E. Ward. 1995. Risk-Pooling Along a Fixed Delivery Route Using a Dynamic Inventory-allocation Policy. *Management Science* **41**, 344-362.

Kushner, H. J. 1977. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations.* Academic Press, New York, NY.

Kushner, H.J. and P.G. Dupuis. 1992. *Numerical Methods for Stochastic Control Problems in Continuous Time.* Springer-Verlag. New York, NY.

Larson, R. C. 1988. Transporting Sludge to the 106-Mile Site: An Inventory/Routing Model for Fleet Sizing and Logistics System Design. *Transportation Science* **22**, 186-198.

Markowitz, D. M., M. I. Reiman and L. M. Wein. 1995. The Stochastic Economic Lot Scheduling Problem: Heavy Traffic Analysis of Dynamic Cyclic Policies. Sloan School of Management, MIT, Cambridge, MA.

Minkoff, A. S. 1993. A Markov Decision Model and Decomposition Heuristic for Dynamic Vehicle Dispatching. *Operations Research* **41**, 77-90.

Peterson, W. P. 1991. A Heavy Traffic Limit Theorem for Networks of Queues with Multiple Customer Types. *Mathematics of Operations Research* **16**, 90-118.

Reiman, M. I. 1984. Open Queueing Networks in Heavy Traffic. *Mathematics of Operations Research* **9**, 441-458.

Rubio, R. 1995. Dynamic-Stochastic Vehicle Routing and Inventory Problem. Ph.D. Thesis, Operations Research Center, MIT, Cambridge, MA.

Simchi-Levi, D. 1992. Hierarchical Planning for Probabilistic Distribution Systems in Euclidean Spaces. *Management Science* **38**, 198-211.

Trudeau, P. and M. Dror. 1992. Stochastic Inventory Routing: Route Design with Stockouts and Route Failures. *Transportation Science* **26**, 171-184.

Wein, L.M. 1992. Dynamic Scheduling of a Multiclass Make-to-Stock Queue. *Operations Research*. **40**, 724-735.

Date